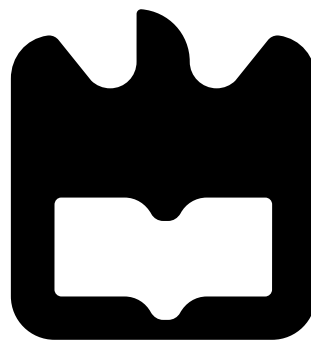




**Tiago André  
da Silva**

**Modelação Acústica de Salas e Auralização em  
Tempo Real**







**Tiago André  
da Silva**

## **Modelação Acústica de Salas e Auralização em Tempo Real**

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia Electrónica e Telecomunicações, realizada sob a orientação científica do Doutor António Guilherme Rocha Campos e do Doutor Paulo Miguel de Jesus Dias, Professores Auxiliares do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro.



**o júri / the jury**

presidente / president

**José Manuel Neto Vieira**

Professor Auxiliar do Dep. de Electrónica, Telecomunicações e Informática da Universidade de Aveiro

vogais / examiners committee

**António Guilherme Rocha Campos**

Professor Auxiliar do Dep. de Electrónica, Telecomunicações e Informática da Universidade de Aveiro (orientador)

**Salviano Filipe Silva Pinto Soares**

Professor Auxiliar do Dep. de Engenharias da Universidade de Trás-os-Montes e Alto Douro



## Resumo

O trabalho apresentado neste documento visa criar modelos de realidade virtual áudio em tempo real, tão convincentes quanto possível, que permitam auralização em visitas virtuais (*walkthrough auralisation*).

São estudados modelos físicos de simulação acústica de salas - a melhor opção do ponto de vista da capacidade de atingir resultados perceptualmente convincentes. É descrita a implementação de um pacote de modelação com base em guias-de-ondas digitais (*DWM*).

Os modelos deste tipo tornam impraticável, pelo seu peso computacional, o cálculo das respostas impulsivas (RIR) em tempo real. Para contornar esta dificuldade, propõe-se que as respostas impulsivas sejam pré-calculadas, em pontos escolhidos ao longo de um trajeto pré-definido dentro da sala, com determinadas orientações da cabeça do ouvinte. Esta abordagem de pré-cálculo impõe certas restrições de movimentação ao utilizador mas, dentro delas, torna possível a auralização em tempo real recorrendo a algoritmos de convolução *Overlap-Add*.

O pré-cálculo de respostas impulsivas implica um grande peso computacional. Para atenuar o problema, explora-se a interpolação espacial entre respostas impulsivas. São discutidos os principais métodos de interpolação propostos na literatura, as suas vantagens e inconvenientes relativos. Evidenciam-se certas dificuldades de aplicação prática, não encontradas na literatura, para as quais são propostas soluções.

São apresentados testes de validação e exemplos demonstrativos dos problemas encontrados e do desempenho das soluções preconizadas.





## Abstract

The work presented in this dissertation aims at creating real-time audio virtual reality (VR) models, as convincing as possible, to allow walkthrough auralisation applications.

It explores physical models for room acoustic simulation - the best option in terms of the ability to provide perceptually convincing results. A software package implemented for digital waveguide mesh modelling is described.

The computational burden of this kind of models makes it impossible to obtain room impulse responses (RIR) in real time. The proposed method of circumventing this difficulty is RIR pre-calculation at selected points along a pre-defined path within the room, with given listener head orientations. This pre-processing approach imposes movement restrictions to the user but, within those restrictions, allows real-time auralisation using the Overlap-Add convolution algorithm.

RIR pre-calculation involves significant computational burden. To mitigate the problem, spatial RIR interpolation is explored. The main interpolation methods proposed in the literature, with their relative advantages and disadvantages, are discussed. One of them is investigated in greater detail through a case study.

Certain practical application difficulties are exposed and solutions are proposed for them.

Validation tests are presented, along with demonstrative examples of the problems found and the performance of the solution proposed.



# Conteúdo

<b>Conteúdo</b>	<b>i</b>
<b>Lista de Figuras</b>	<b>iii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Motivação . . . . .	1
1.2 Objetivos . . . . .	2
1.3 Organização . . . . .	2
<b>2 Modelação Acústica</b>	<b>5</b>
2.1 Conceitos-Chave . . . . .	5
2.2 Técnicas de Modelação . . . . .	7
2.3 Modelação Analógica . . . . .	8
2.4 Modelação Digital . . . . .	9
2.4.1 Modelos Geométricos . . . . .	10
2.4.2 Modelos Físicos . . . . .	10
<b>3 Modelação por Guias de Onda Digitais</b>	<b>13</b>
3.1 Introdução . . . . .	13
3.2 Princípio de Funcionamento . . . . .	16
3.3 <i>Software</i> de Modelação . . . . .	18
3.3.1 Estrutura Geral . . . . .	18
3.3.2 Configuração da Simulação . . . . .	18
3.3.3 Caracterização Acústica dos Materiais . . . . .	19
3.3.4 Resultados da Simulação . . . . .	19
3.3.5 Algoritmo de Simulação . . . . .	20
3.3.6 Validação . . . . .	21
<b>4 Auralização em Tempo Real</b>	<b>23</b>
4.1 Introdução . . . . .	23
4.2 Convolução pelo Método <i>Overlap-Add</i> . . . . .	23
4.3 Transição entre RIR e Interpolação Temporal . . . . .	27
4.4 Implementação Prática . . . . .	27
4.5 Auralização num Trajeto Virtual . . . . .	30

<b>5</b>	<b>Interpolação de Respostas Impulsionais</b>	<b>35</b>
5.1	Introdução . . . . .	35
5.2	Interpolação Linear . . . . .	35
5.3	Interpolação em Fase e Módulo . . . . .	38
5.4	Aplicação de Interpolação em Fase e Módulo . . . . .	41
5.5	Uma Questão de Energia . . . . .	42
5.6	Testes . . . . .	44
<b>6</b>	<b>Conclusões e Trabalho Futuro</b>	<b>45</b>
6.1	Conclusões . . . . .	45
6.2	Trabalho Futuro . . . . .	45
6.2.1	Perspectiva Geral . . . . .	45
6.2.2	Modelação Acústica . . . . .	46
6.2.3	Auralização em Tempo Real . . . . .	47
<b>A</b>	<b>Caraterísticas Acústicas de Materiais</b>	<b>48</b>
A.1	Impedância Acústica Específica . . . . .	48
A.2	Coeficiente de Reflexão . . . . .	50
<b>B</b>	<b>Código-Fonte do Software de Modelação <i>DWsim</i></b>	<b>52</b>
<b>C</b>	<b>Código-Fonte da Biblioteca <i>OLAlib</i></b>	<b>53</b>
<b>D</b>	<b>Auralização num Trajeto Virtual</b>	<b>54</b>
<b>E</b>	<b>Interpolação de Sinusóides</b>	<b>55</b>
<b>F</b>	<b>Interpolação em Módulo e Fase de uma RIR</b>	<b>59</b>
<b>G</b>	<b>Resultados dos Testes de Interpolação</b>	<b>60</b>
	<b>Bibliografia</b>	<b>63</b>

# Lista de Figuras

1.1	Processo de modelação e auralização de uma sala . . . . .	2
2.1	Fenómenos de reflexão, absorção e refração de uma onda sonora . . . . .	6
2.2	Diagrama de blocos de um sistema bi-aural . . . . .	8
2.3	Técnicas de modelação acústica . . . . .	9
3.1	Guia de onda digital num meio de propagação de impedância acústica $Z_0$ . . .	13
3.2	Processo de excitação e captação de uma guia de onda digital . . . . .	14
3.3	Guia de onda digital na transição entre meios de propagação com impedância acústica distinta. . . . .	15
3.4	Topologias de malhas de guias de onda digitais [1] . . . . .	16
3.5	Técnicas de modelação acústica . . . . .	18
3.6	Representação tridimensional de um nó da malha. . . . .	21
3.7	Modos de ressonância teóricos e simulados. . . . .	22
4.1	Esquema de funcionamento do método <i>Overlap-Add</i> [4] . . . . .	26
4.2	Estrutura de dados <i>OlaStrct</i> . . . . .	28
4.3	Pseudo-código da implementação do método <i>OLA</i> utilizando a biblioteca <i>OLALib</i> . . . . .	29
4.4	Representação tridimensional da sala utilizada no estudo de caso . . . . .	30
4.5	Trajeto auralizável na sala de reuniões . . . . .	31
4.6	Representação da cabeça do ouvinte por uma matriz tridimensional de pontos. . . . .	32
4.7	Aplicação (em <i>Matlab</i> ) que permite auralizar em tempo real a sala considerada no estudo de caso. . . . .	34
5.1	Esquema de interpolação linear de respostas impulsionais. . . . .	36
5.2	Interpolação linear entre duas sinusóides desfasadas com frequência e amplitude iguais. . . . .	37
5.3	Esquema de interpolação de respostas impulsionais na presença da fonte sonora. . . . .	38
5.4	Interpolação em módulo e fase entre duas sinusóides desfasadas com amplitude e frequência iguais. . . . .	41
5.5	Módulo das sinusóides interpoladas e da sinusóide resultante da interpolação . . . . .	42
5.6	Resposta impulsional sintetizada por interpolação em fase e módulo com <i>unwrap</i> de fases . . . . .	43
5.7	Esquema de teste para interpolação de respostas impulsionais. . . . .	44
E.1	Resultados da interpolação entre sinusóides de igual amplitude . . . . .	56
E.2	Resultados da interpolação entre sinusóides em que uma tem o dobro da amplitude . . . . .	57

### E.3 Resultados da interpolação entre sinusóides em que uma tem o triplo da amplitude 58

# Capítulo 1

## Introdução

### 1.1 Motivação

A forma como o ser humano percebe o som não depende unicamente da fonte sonora mas também do meio em que o som se propaga até ser captado pelos seus ouvidos.

A propagação das ondas sonoras e a maneira como estas interagem fisicamente com o meio ambiente é o principal objecto de estudo da Acústica. A percepção cognitiva dos sinais sonoros captados pelos ouvidos é tema de estudo de uma outra área científica directamente relacionada: a Psicoacústica.

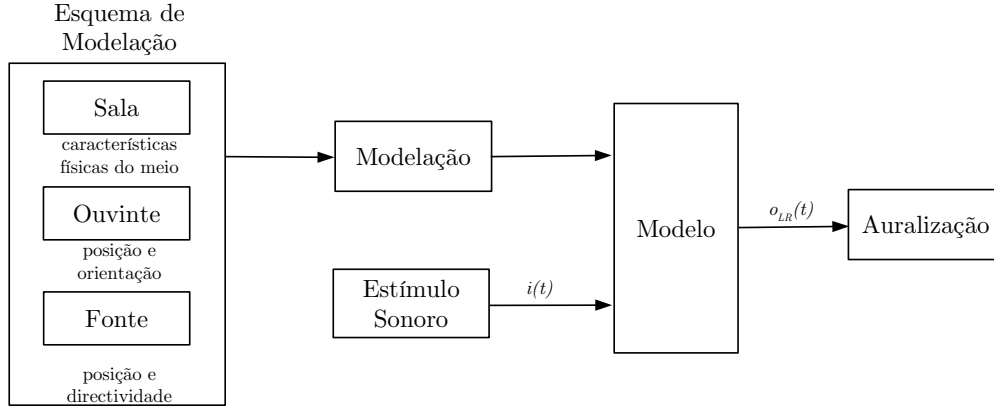
A Acústica Arquitectural é o ramo da acústica que estuda o comportamento sonoro em edifícios. Compreende essencialmente duas áreas de estudo: isolamento sonoro (comportamento face a fontes sonoras externas indesejáveis, que constituem ruído) e acústica de salas (comportamento face a fontes sonoras internas cuja audição se deseja otimizar).

A Acústica de Salas estuda o comportamento do som em espaços fechados ou semiabertos. Tem em conta as características físicas dos materiais que continuam o espaço e fatores psicológicos do ouvinte, pelo que a Psicoacústica desempenha neste campo um papel muito importante, embora não abordado neste trabalho.

Do ponto de vista acústico, uma sala pode ser considerada um sistema. O seu comportamento quando sujeita a um determinado estímulo - no caso, um estímulo sonoro - pode ser descrito por equações matemáticas - no caso, a 'equação de onda' de Helmholtz. A sua resolução para um espaço fisicamente delimitado requer o estabelecimento de condições-fronteira adequadas, para traduzir os fenómenos (reflexão, absorção, difração, refração, interferência...) devidos à interacção das ondas com os obstáculos presentes (paredes, chão, teto, mobiliário...) e eventual heterogeneidade (em termos de temperatura e humidade, por exemplo) do meio no interior da própria sala. Torna-se por isso impraticável uma resolução analítica, exceto em situações ideais (i.e. extremamente simplificadas). A resposta a esta dificuldade é dada pelo desenvolvimento de modelos acústicos.

Os modelos acústicos permitem definir as condições que influenciam a percepção sonora (sinal sonoro emitido pela(s) fonte(s), sua diretividade e posição na sala; posição e orientação da cabeça do ouvinte; características dos materiais de revestimento e mobiliário da sala) e recriar os estímulos sonoros que atingem os ouvidos nessas condições: a este processo chama-se 'Auralização'.

Na figura 1.1 podemos observar que o processo de modelação requer essencialmente dois tipos de dados: informação sobre a representação física da sala (incluindo ouvinte e fonte) e



*Figura 1.1: Processo de modelação e auralização de uma sala*

o estímulo sonoro (voz, instrumento, gravação...) que se pretende auralizar,  $i(t)$ . O modelo resultante permite obter, para um dado sinal de entrada  $i(t)$ , o sinal de saída,  $O_{LR}(t)$ , composto pelos estímulos sonoros no ouvido esquerdo (L) e no ouvido direito (R) nas condições apresentadas pelo esquema de modelação (sala, ouvinte, fonte). Torna-se assim possível a auralização, com recurso a auscultadores.

A auralização em tempo real pode ser definida como a capacidade de obter e reproduzir o sinal de saída à medida que o sinal de entrada é recebido. Constitui um grande desafio tecnológico. Sublinhe-se que a presença do ouvinte na sala afeta a sua percepção auditiva e por isso pode ser necessário obter dados para auralização em várias posições do ouvinte e eventualmente também para várias posições da(s) fonte(s). Dependendo do objetivo pretendido, isso pode exigir múltiplos processos de modelação e, por consequência, múltiplos modelos.

## 1.2 Objetivos

O trabalho realizado no âmbito desta dissertação tem como objetivo desenvolver ferramentas que possibilitem modelar acusticamente uma sala e auralizá-la em tempo real. Para isso é necessário começar por estudar e compreender as técnicas de modelação e auralização apresentadas na literatura e analisar quais e como devem ser aplicadas.

## 1.3 Organização

A estrutura deste documento segue essencialmente a ordem cronológica do trabalho realizado. À presente Introdução seguem-se cinco capítulos:

**Capítulo 2** (Modelação Acústica)- Apresenta os fundamentos teóricos da modelação acústica e estuda as técnicas de modelação presentes na literatura. São analisadas as suas limitações e a forma como estas podem condicionar a concretização dos objetivos propostos.



**Capítulo 3** (Modelação por Guias de Onda Digitais)- Explica o princípio de funcionamento da técnica de modelação adoptada. Trata-se de modelação por malha de guias de onda digitais - *Digital Waveguide Mesh (DWM) modelling*. É apresentado o pacote de software em C++ desenvolvido para implementar esta técnica, bem como os testes de validação efectuados.

**Capítulo 4** (Auralização em Tempo Real)- Explora técnicas computacionais para auralização em tempo real. É apresentado um estudo de caso aplicando as técnicas de modelação e auralização implementadas.

**Capítulo 5** (Interpolação de Respostas Impulsionais)- Centra-se no problema da interpolação espacial de respostas impulsionais, começando pelo estudo da literatura que aborda este tema. São apresentados os problemas concretos que resultam dos diferentes processos de interpolação possíveis. É explicado o processo de desenvolvimento da técnica de interpolação implementada e apresentados os resultados obtidos.

**Capítulo 6** (Conclusão e Trabalho Futuro)- São apresentadas as conclusões finais sobre o trabalho realizado e é proposto trabalho futuro no sentido de melhorar as técnicas implementadas.

Ulteriormente apresentam-se os Apêndices A, B, C, D, E, F e G que terminam com a enumeração da Bibliografia utilizada durante a pesquisa deste trabalho de dissertação.



## Capítulo 2

# Modelação Acústica

### 2.1 Conceitos-Chave

A propagação do som no espaço a três dimensões é um fenómeno ondulatório descrito matematicamente pela equação de Helmholtz:

$$\frac{\partial^2}{\partial x^2}p(t, x, y, z) + \frac{\partial^2}{\partial y^2}p(t, x, y, z) + \frac{\partial^2}{\partial z^2}p(t, x, y, z) = \frac{1}{c^2} \frac{\partial^2}{\partial t^2}p(t, x, y, z) \quad (2.1)$$

Nesta equação,  $p$  representa a variação de pressão num determinado ponto  $(x, y, z)$  do espaço ao longo do tempo  $t$ , e  $c$  é a velocidade de propagação do som. Esta é, aproximadamente, constante para um determinado meio e depende das características físicas dos materiais que o constituem. Pode ser definida por:

$$c = \frac{z}{\rho} \quad (2.2)$$

em que  $z$  representa a impedância acústica característica do meio, e  $\rho$  a sua massa volúmica. No anexo A é apresentada uma tabela de impedâncias acústicas características para diferentes materiais.

Na interação de uma onda sonora com o meio físico podem ocorrer vários fenómenos bem conhecidos:

- **Reflexão** É originado pela transição de uma onda sonora entre meios físicos com impedâncias acústicas distintas. Considere-se o exemplo ilustrado na figura 2.1a. Quando atinge a fronteira que separa os meios A e B, a onda proveniente do meio A (onda incidente) pode ser (pelo menos parcialmente) refletida de volta ao meio A. A velocidade e a frequência da onda não sofrem alteração na componente refletida. No entanto, a sua amplitude é reduzida por um fator  $r$ , designado por coeficiente de reflexão. Este depende das características físicas dos materiais que constituem os meios de propagação, bem como da frequência e ângulo de incidência da onda. Quando a superfície que separa os dois meios é aproximadamente plana, o fenómeno é

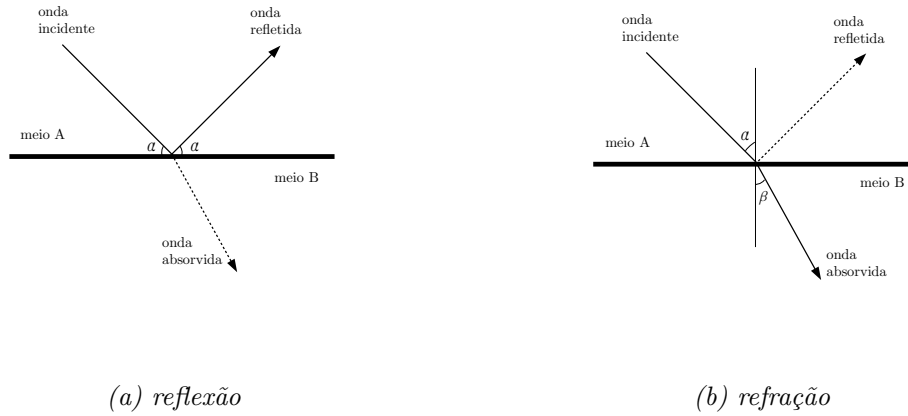


Figura 2.1: Fenómenos de reflexão, absorção e refração de uma onda sonora

descrito pela lei de Snell para a reflexão: as ondas incidente e reflectida formam ângulos iguais com a superfície.

No anexo A é apresentada uma tabela onde se podem encontrar coeficientes de reflexão para diferentes materiais (meios) de propagação.

- **Absorção** Dois tipos podem ser distinguidos:

1. Absorção na transição entre meios de diferentes impedâncias. De forma complementar ao fenómeno de reflexão, a onda incidente é (pelo menos parcialmente) absorvida pelo novo meio de propagação.
2. Absorção do próprio meio por dissipação de energia sob forma térmica. O meio de propagação apresenta propriedades de absorção directamente relacionadas com a temperatura e humidade.

- **Refração** Manifesta-se de duas maneiras:

1. Mudança repentina de direcção originada pela transição entre meios de propagação com impedâncias acústicas características diferentes. Este fenómeno está relacionado com o fenómeno de absorção, como mostra a figura 2.1b. A direcção da onda absorvida (que sofre refração) pode ser obtida recorrendo à lei de Snell para a refração,

$$\frac{\sin(\alpha)}{c_A} = \frac{\sin(\beta)}{c_B} \quad (2.3)$$

onde  $\alpha$  e  $\beta$  são os ângulos das ondas incidente e refratada com a normal ao plano de incidência e  $c_A$  e  $c_B$  são as respectivas velocidades de propagação (meio A e meio B).

2. Propagação curvada devido à heterogeneidade do meio. Por exemplo, a existência de gradientes de temperatura no meio altera a velocidade de propagação da onda mudando a sua direção gradualmente ao longo do espaço.

- **Difração** Ocorre quando uma onda sonora contorna ou transpõe um obstáculo encontrado no seu caminho. Este fenómeno depende da relação entre o comprimento de onda e a dimensão do obstáculo em causa.
- **Espalhamento e Difusão** Quando uma superfície irregular é atingida por uma onda sonora, a combinação dos fenómenos de reflexão e difração resulta no espalhamento de várias ondas, em várias direcções. A difusão é a situação limite em que a superfície é tão irregular que as ondas sonoras são espalhadas em praticamente todas as direcções.
- **Interferência** Quando várias ondas sonoras partilham o mesmo espaço de propagação, os diversos pontos estão sujeitos à sua influência conjunta. Em cada ponto do espaço, a pressão sonora resulta da sobreposição das pressões sonoras criadas pelas várias ondas. A interferência diz-se 'construtiva' quando as componentes de pressão se somam em valor absoluto e 'destrutiva' quando se subtraem em valor absoluto (podendo, no limite, anular-se).

Todos estes fenómenos são intrinsecamente considerados pela equação de *Helmholtz*. É necessário, no entanto, estabelecer adequadas condições-fronteira, tendo em conta a forma e constituição física dos limites do espaço. Isso é impraticável de forma analítica. É, por isso, necessário recorrer a modelos.

## 2.2 Técnicas de Modelação

Em *Teoria dos Sistemas*, resposta impulsional é o sinal de saída de um sistema quando é aplicado na sua entrada um sinal de energia infinita durante um intervalo de tempo infinitesimal, ou seja, um impulso de *Dirac*. A resposta impulsional de um sistema linear e invariante no tempo (linear, time-invariant - LTI) caracteriza completamente o comportamento do sistema permitindo obter o sinal de saída por convolução com o sinal de entrada.

A equação de onda de *Helmholtz* é uma equação diferencial com coeficientes constantes (considerando a velocidade de propagação constante para um determinado meio) descrevendo assim um sistema LTI. Isto significa que uma sala, do ponto de vista acústico, pode ser considerada um sistema LTI, pelo que a sua resposta impulsional caracteriza o seu comportamento sonoro completamente.

A resposta impulsional acústica num dado ponto de uma sala é designada na literatura por RIR (*room impulse response*). Pode ser medida directamente, injetando um impulso (*Dirac*) sonoro num outro ponto da mesma sala, ou indirectamente, injetando um sinal de teste (tipicamente um *chirp* logarítmico) e aplicando adequadas técnicas de pós-processamento da resposta recolhida.

A auralização de uma sala, para dada posição da(s) fonte(s), pode ser conseguida recorrendo a uma BRIR (*bi-aural room impulse response*). Uma BRIR é formada por um par de

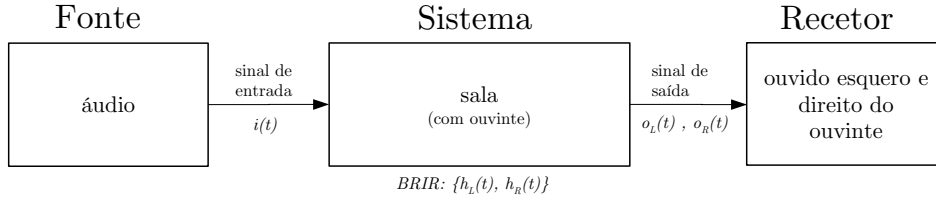


Figura 2.2: Diagrama de blocos de um sistema bi-aural

respostas impulsivas (RIR) referentes a cada ouvido (esquerdo e direito) e caracteriza acusticamente a sala para uma determinada posição e orientação da cabeça do ouvinte. É necessário que o ouvinte integre o modelo da sala, pois a sua presença afeta o campo sonoro no seu interior. Note-se que o próprio pavilhão auditivo do ouvinte tem um papel preponderante na sua percepção auditiva.

A figura 2.2 apresenta um diagrama de blocos de um sistema acústico para uma determinada posição do ouvinte e da(s) fonte(s) sonora(s) numa sala. A BRIR do sistema compreende as respostas impulsivas no ouvido esquerdo,  $h_L(t)$ , e direito,  $h_R(t)$ . O sinal de saída é um sinal estéreo composto por  $o_L(t)$  e  $o_R(t)$ , obtido convolucionando o sinal de entrada,  $i(t)$ , com as respostas impulsivas  $h_L(t)$  e  $h_R(t)$ . O sinal de saída representa os estímulos auditivos que seriam escutados pelo ouvinte no interior da sala se o sinal de áudio  $i(t)$  fosse emitido na posição correspondente à fonte sonora.

O sinal de entrada deve ser desprovido de qualquer registo de interação com o meio onde foi gravado, de forma a representar exclusivamente a fonte sonora. Por isso, a sua gravação é, normalmente, efetuada em câmara anecóica.

A resposta impulsional constitui uma ferramenta essencial; obtê-la é o grande objectivo das técnicas de modelação acústica. Estas dividem-se essencialmente em duas categorias: modelação digital e modelação analógica, como mostra a figura 2.3.

## 2.3 Modelação Analógica

Modelos analógicos foram usados durante muito tempo para sistemas acústicos, nomeadamente salas.

Uma das técnicas utilizadas recorre à observação da propagação de ondas de água num espaço físico delimitado para, por analogia, avaliar o comportamento de ondas sonoras. A sua aplicação é no entanto restrita a duas dimensões.

O uso de fontes de luz (em particular *laser*) permite também simular o comportamento

Modelação Análogica	Modelos baseados em raios de luz (“lasers”)	
	Modelos baseados em ondas aquáticas	
	Modelos baseados em ultrassons	
Modelação Digital	Modelos Geométricos	Método “Ray-tracing”
		Método “Image-Source”
	Modelos Físicos	Método das Diferenças Finitas no Domínio do Tempo
		Método dos Elementos Finitos
		Método dos Elementos Fronteira

*Figura 2.3: Técnicas de modelação acústica*

de ondas sonoras (nomeadamente em termos de reflexões) em modelos de escala reduzida. No entanto, a analogia é limitada às altas frequências. Note-se que, devido à velocidade da luz ser praticamente infinita em comparação com a velocidade do som, nenhuma informação temporal pode ser obtida recorrendo a este tipo de modelação.

Das técnicas analógicas conhecidas, apenas a utilização de ultrassons em modelos de escala reduzida encontra actualmente aplicação prática. O princípio é muito simples: a frequência das ondas sonoras é aumentada pelo mesmo fator de escala de redução do modelo. Assim, a relação entre os comprimentos de onda das componentes do som e as dimensões do modelo mantém-se constante. O meio de propagação do modelo deve ser tal que a velocidade de propagação seja mantida em relação às condições reais. Os materiais utilizados no modelo devem ter características de absorção sonora correspondentes aos materiais reais, o que implica uma certa dificuldade na construção do modelo.

Além disso, a modelação baseada em ultrassons requer equipamento específico (miniaturizado) para injeção e captação de ondas no modelo.

Em resumo, os modelos analógicos oferecem várias dificuldades de aplicação prática. Por isso, a sua utilização é cada vez menos frequente face à emergência de modelos digitais cada vez mais acessíveis e eficientes.

## 2.4 Modelação Digital

A modelação digital consiste na representação e simulação computacional do comportamento acústico de espaços físicos, em particular, salas. Os modelos são completamente virtuais - ao contrário dos sistemas analógicos, não requerem a existência de qualquer elemento físico respeitante à sala a modelar.

A modelação física e a modelação geométrica, são as duas grandes vertentes da modelação digital. Ambas possuem vantagens e desvantagens, pelo que a escolha deve ter em conta os

objetivos pretendidos.

### 2.4.1 Modelos Geométricos

Os modelos geométricos constituem uma descrição simplificada da propagação do som baseada no comportamento de ondas óticas. As ondas sonoras são tratadas como raios cuja interação com o meio é descrita pela lei de *Snell*. Como consequência desta aproximação, apenas são considerados os fenómenos de absorção e reflexão. Podem ser obtidos essencialmente recorrendo a duas técnicas:

- **ray-tracing** Esta técnica consiste em seguir o trajeto de um conjunto de raios sonoros (tratados como raios de luz) a partir da fonte até atingir um recetor. As reflexões são descritas pela lei de *Snell*.
- **mirror-image source (MIS)** As reflexões são processadas considerando que cada raio refletido é gerado por uma fonte virtual, que é uma imagem espelhada, relativamente à superfície de reflexão da fonte que emite o raio sonoro incidente nessa superfície de reflexão.

A utilização de modelos geométricos exige que a sala modelada (e objectos internos) seja aproximada por superfícies planas de dimensões muito grandes comparativamente com o comprimento de onda das ondas sonoras.

A gama de frequências audíveis situa-se sensivelmente entre os  $20Hz$  e os  $20KHz$ . Isto significa que (à temperatura ambiente), o comprimento de onda das ondas sonoras audíveis está compreendido entre os  $0.017m$  e os  $17m$ . Facilmente se verifica que os modelos geométricos apenas têm aplicação prática a altas frequências uma vez que, para baixas frequências, as dimensões das superfícies devem ser tão elevadas que impossibilitam representar, de forma aceitável, a forma da sala.

A aplicabilidade destes modelos é, por isso, dependente da geometria da sala a modelar e do detalhe pretendido no processo de modelação.

Os modelos geométricos podem ser facilmente combinados com filtros *head-related transfer functions (HRTF)*, permitindo assim auralização bi-aural. Cada *HRTF* caracteriza a forma como o ouvido humano capta o som emitido por uma fonte sonora emitindo numa determinada direção relativamente à cabeça. É obtida experimentalmente colocando pequenos microfones nos ouvidos de um manequim, e registando o som de chegada para a respectiva direção da cabeça em relação à fonte sonora. As gravações de *HRTF* são realizadas de forma anecóica.

As limitações dos modelos geométricos são, até certo ponto, compensadas pela sua carga computacional relativamente baixa. A auralização de uma sala pode ser obtida sem grande dificuldade, no entanto, os modelos constituem aproximações pouco rigorosas da realidade, sendo os resultados obtidos satisfatórios apenas em determinadas condições (salas de forma regular delimitadas por superfícies planas, bandas de frequência elevadas).

### 2.4.2 Modelos Físicos

Os modelos físicos baseiam-se na resolução da equação de onda de *Helmholtz* (sendo por isso também conhecidos como *wave-based*). Constituem aproximações numéricas da sua solução, permitindo assim simular o comportamento do campo sonoro de uma sala com determinadas



caraterísticas (que pode ou não corresponder a uma sala existente na realidade). Implicam discretização do espaço e do tempo, uma vez que envolvem processos iterativos que simulam o comportamento das partículas que compõem o meio.

Dependendo da resolução da discretização e das dimensões reais da sala, os modelos físicos podem ser, do ponto de vista computacional, extremamente pesados, exigindo longos períodos de simulação. Por outro lado, uma elevada resolução de discretização permite descrever a sala com grande pormenor, independentemente da sua forma e das fronteiras. Isto torna-se uma grande vantagem face aos modelos geométricos. Os modelos físicos são, por isso, modelos potencialmente muito rigorosos.

Existem vários métodos de modelação física. Os mais utilizados em Acústica são os seguintes:

- **Método dos Elementos Finitos** (*Finite Element Method - FEM*)
- **Método dos Elementos Fronteira** (*Boundary Element Method - BEM*)
- **Método das Diferenças Finitas no Domínio do Tempo** (*Finite Difference Time Domain - FDTD*)

A modelação por guias-de-onda digitais (*Digital Waveguide - DW*), é uma técnica que se insere no método FDTD. Baseia-se na aproximação numérica da solução de d'Alembert para a equação de onda a uma dimensão (1-D). Foi desenvolvida, originalmente, para aplicações musicais, nomeadamente vibração de cordas, obtendo-se resultados de sons sintetizados muito satisfatórios comparativamente com os sons reais reproduzidos nas mesmas condições físicas.

Uma guia de onda digital pode ser interligada com outras guias de onda formando assim, uma malha a duas dimensões (2-D) denominada na literatura por *Digital Waveguide Mesh (DWM)*. Malhas de guias de onda digitais possibilitam simular a propagação de onda em duas dimensões espaciais e assim sintetizar sons produzidos por superfícies acústicas (e.g. tambores) [2]. A modelação *DWM* a três dimensões surge como uma extensão natural, directamente aplicável em Acústica de Salas.

Pretende-se neste trabalho desenvolver ferramentas que permitam modelar uma determinada sala com um grau de rigor tão elevado quanto possível. Daí a opção pela utilização de modelos físicos.

O capítulo seguinte estuda a modelação física por malhas de guias-de-onda digitais *DWM* tri-dimensionais.



## Capítulo 3

# Modelação por Guias de Onda Digitais

### 3.1 Introdução

Os modelos físicos permitem descrever acusticamente um espaço físico considerando todos os fenómenos ondulatórios estudados no capítulo anterior (não apenas reflexão e absorção, como nos modelos geométricos, mas também, nomeadamente, difração e interferência). A modelação por guias de onda digitais faz parte desta categoria.

Uma guia de onda digital (*Digital Waveguide* - *DW*) resulta da discretização espacial e temporal da solução de d'Alembert da equação de onda unidimensional. Esta consiste na sobreposição de duas ondas que se propagam na mesma direção mas em sentidos opostos. Cada onda pode ser representada por um par de linhas de atrasos digitais unitários. Um atraso digital de  $N$  amostras é descrito, no domínio do tempo discreto, pela função  $\delta(n - N)$ . A sua transformada em frequência (para sinais discretos)  $Z$  é dada por:

$$\delta(n - N) \xrightarrow{Z} Z^{-N} \quad (3.1)$$

Uma linha de atrasos pode, portanto, ser representada por um filtro digital com função transferência  $Z^{-N}$ , sendo uma guia de onda digital constituída por um par de filtros digitais, como ilustra a figura 3.1.

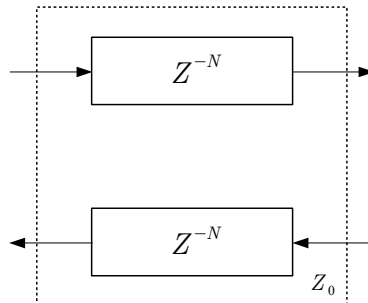
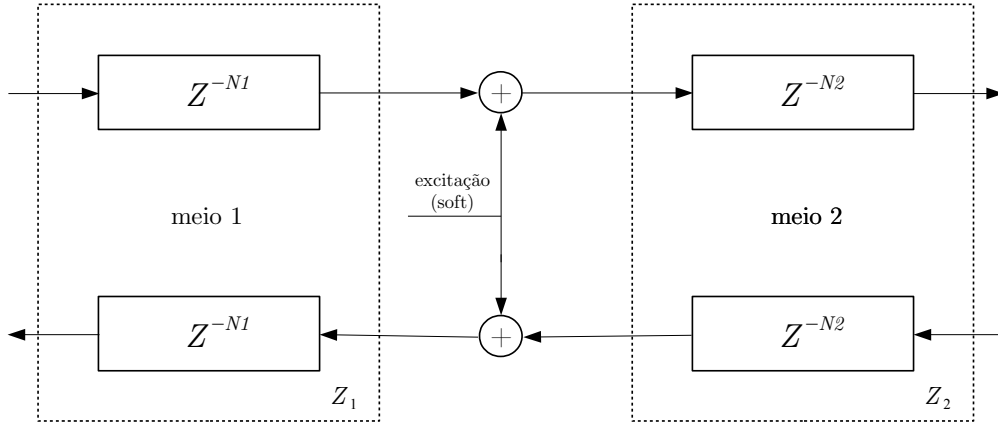
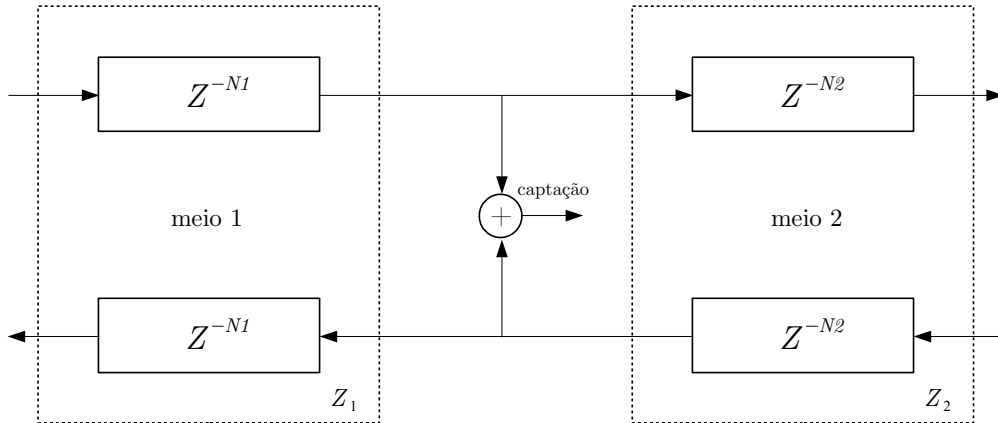


Figura 3.1: Guia de onda digital num meio de propagação de impedância acústica  $Z_0$ .



(a) *Excitação da guia de onda digital*



(b) *Captação da guia de onda digital*

Figura 3.2: Processo de excitação e captação de uma guia de onda digital

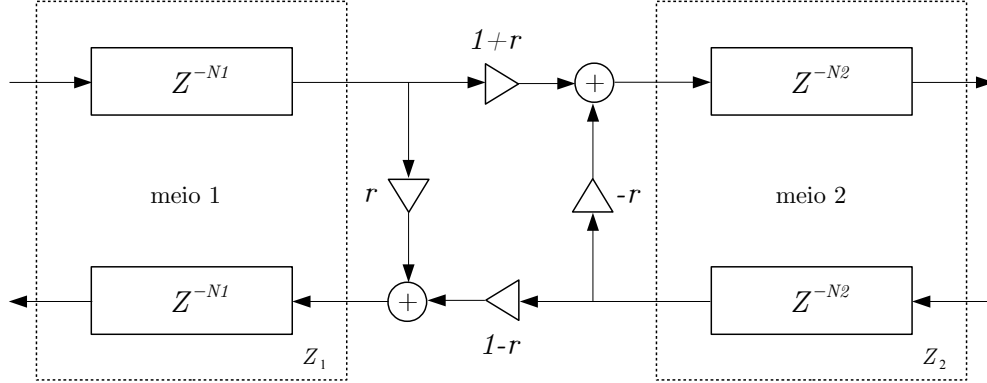


Figura 3.3: Guia de onda digital na transição entre meios de propagação com impedância acústica distinta.

A descrição do comportamento de uma onda sonora depende do meio onde se propaga. Isto implica que cada guia de onda digital representa uma onda sonora num determinado meio de propagação, caracterizado por uma impedância acústica  $Z$  (não confundir com transformada  $Z$ ). Na figura 3.1, o meio de propagação é caracterizado por uma impedância acústica  $Z_0$ .

Os valores de pressão sonora são obtidos somando os valores das duas linhas de atraso em qualquer posição. Da mesma forma, é possível excitar a guia de onda digital em qualquer posição da linha de atrasos. A figura 3.2 ilustra estes processos.

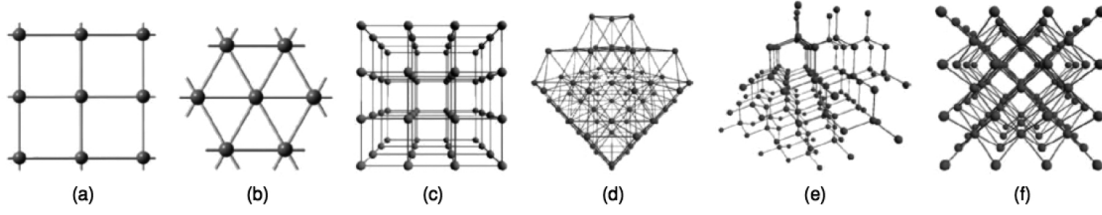
A transição entre guias de onda de diferentes impedâncias é ilustrada na figura 3.3.  $Z_1$  e  $Z_2$  representam as impedâncias acústicas dos meios de propagação 1 e 2, respetivamente. O coeficiente de reflexão  $r$ , é dado em por:

$$r = \frac{Z_2 - Z_1}{Z_2 + Z_1} \quad (3.2)$$

Uma malha de guias de onda digitais (*digital waveguide mesh*) pode ser obtida interligando várias linhas de atraso. Os pontos onde estas se interligam são denominados nós (*scattering junctions*). Uma malha bidimensional (2-D) de guias de ondas digitais permite modelar comportamentos acústicos de superfícies planas vibrantes. A sua expansão para três dimensões (3-D) pode ser aplicada a espaços físicos tridimensionais possibilitando modelar comportamentos acústicos no interior destes. A expansão da guia de onda digital para malhas 2 ou 3-D, implica naturalmente, mais recursos computacionais.

Um espaço pode ser representado por uma malha de nós em que cada nó está conetado a nós vizinhos por linhas de atraso bidirecionais. O número de conexões associadas a cada nó depende do número de nós vizinhos. A distribuição geométrica dos nós caracteriza a topologia da malha de guias de onda digitais. A figura 3.4 ilustra algumas das topologias mais utilizadas neste tipo de modelação.

Neste capítulo, é estudado o princípio de funcionamento da modelação por guias de onda



*Figura 3.4: Topologias de malhas de guias de onda digitais [1]*  
 (a) retilínea 2-D; (b) triangular 2-D; (c) retilínea 3-D; (d) dodecaédrica 3-D; (e) tetraédrica 3-D; (f) octaédrica 3-D

digitais e as suas limitações. Por fim, é apresentado um software que foi concebido para modelar salas e explicado o seu funcionamento.

## 3.2 Princípio de Funcionamento

Como já referido, um nó é um ponto onde se conectam guias de onda digitais. Cada guia de onda digital propaga-se num meio caracterizado por uma impedância acústica e contribui com uma determinada pressão no ponto de conexão (nó).

Cada nó obedece a duas simples condições:

- a soma das velocidades das ondas que convergem no nó (incidentes) deve ser igual à soma das velocidades das ondas que divergem do nó (reflectidas).

$$\sum_{i=1}^n v_i^+ = \sum_{i=1}^n v_i^- \quad (3.3)$$

- a pressão em cada guia de onda digital ( $p_i = p_i^+ + p_i^-$ ) conectada a um nó é igual à pressão nesse mesmo nó .

$$p_1 = p_2 = p_3 = \dots p_n = p \quad (3.4)$$

A pressão num determinado nó (sem perdas),  $p$ , pode ser obtida em função da pressão das ondas incidentes,  $p_i^+$  e em função da impedância acústica do meio de propagação ( $Z_i = \frac{p_i}{v_i}$ ) pela seguinte expressão:

$$p = \frac{2 \cdot \sum_{i=1}^n \frac{p_i^+}{Z_i}}{\sum_{i=1}^n \frac{1}{Z_i}} \quad (3.5)$$

A excitação sonora da malha pode ser realizada fazendo incidir ondas de pressão em

qualquer nó da malha. Suponha-se que se pretende injetar (num determinado instante) uma pressão  $p_n$  num determinado nó  $n$ . O valor de pressão,  $p_i$ , de cada onda incidente nesse nó vem dado por:

$$p_i^+ = \frac{p_n}{2} \quad (3.6)$$

Se a onda incidente injetada no nó for somada à onda incidente que eventualmente se propaga nesse instante, a excitação da fonte diz-se suave (*soft*). Se a onda incidente que se propaga nesse nó for descartada e apenas for considerada a onda incidente injetada, a excitação diz-se dura (*hard*). A medição da pressão sonora é obtida medindo diretamente a pressão de um determinado nó.

Considere-se a topologia retilínea tridimensional em que cada nó está a uma distância  $d$  dos nós vizinhos. Fazendo corresponder a cada nó da malha o tipo de material que ocupa a mesma posição relativa, é possível representar qualquer espaço físico por uma malha deste tipo. A representação de espaços com grande nível de detalhe requer uma densidade de pontos maior. A distância internodal (entre nós),  $d$ , é obtida por:

$$d = \frac{c \cdot \sqrt{3}}{fa} \quad (3.7)$$

onde  $c$  representa a velocidade do som e  $fa$  a frequência de amostragem do sinal injectado no modelo. Isto significa que as dimensões físicas reais são definidas pela frequência de amostragem do sinal áudio que se pretende auralizar.

O modelo pode não representar as dimensões exactas do espaço físico modelado. Se uma distância real não for múltipla de  $d$ , isso implica uma aproximação. Na modelação de edifícios ou salas, esta aproximação pode ser desprezada. Repare-se que, pela equação 3.7, a uma frequência de amostragem  $fa = 44100Hz$ ,  $d = 1.35cm$ . Trata-se, portanto, de um valor muito pequeno comparado com as dimensões médias de uma sala comum. A escolha da distância internodal,  $d$ , no processo de modelação, deve ter em conta fatores como o rigor da representação da sala e peso computacional. Quanto maior o número de nós da malha, maior a definição física do modelo e mais pesado este se torna do ponto de vista computacional.

Outro fator importante nesta matéria é o erro de dispersão do modelo: a malha de nós (nomeadamente a retilínea) não permite considerar uma velocidade constante de propagação de onda em todas as direções. Este erro depende diretamente da topologia da malha utilizada na modelação e pode ser atenuado à custa do aumento do peso computacional do modelo, ou seja, reduzindo a distância internodal  $d$  (mais nós). No entanto, é uma consequência da discretização espacial, pelo que existe sempre. O erro de dispersão é tema de estudo nesta área e pode ser analisado mais em pormenor em [3].

Na modelação acústica de uma sala, não se pretende obter informação sobre o campo sonoro que se propaga do interior para o exterior da sala (e vice-versa). Ou seja, apenas é considerado o comportamento acústico no interior da sala. Neste sentido podemos considerar que cada nó conetado a um nó de ar é um nó fronteira. Assim, cada nó fronteira pode ser caracterizado por um coeficiente de reflexão  $r$  (equação 3.2) que depende da impedância acústica do ar e do material que o nó representa.

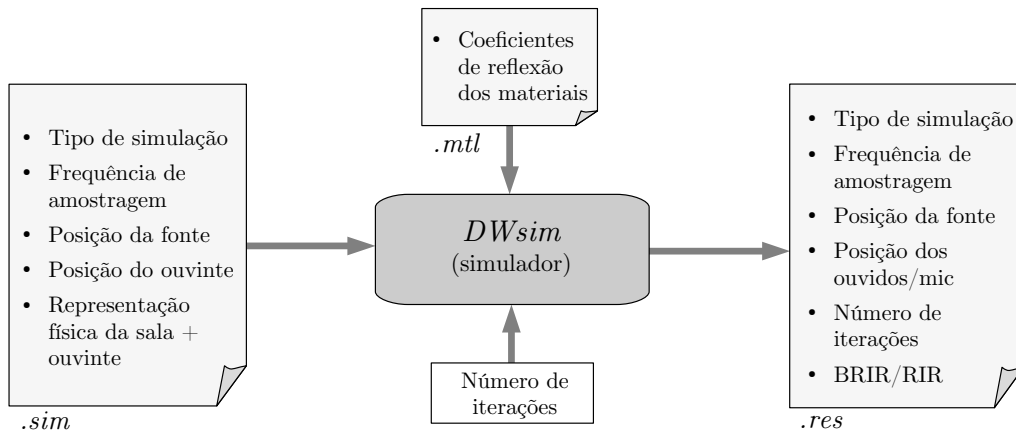


Figura 3.5: Técnicas de modelação acústica

### 3.3 Software de Modelação

#### 3.3.1 Estrutura Geral

Foi desenvolvido um pacote de software (em *C++*), designado *DWsim* (Digital Waveguide Simulator), que permite modelar salas (ou espaços fisicamente delimitados) por malhas de guias de ondas digitais. O *DWsim* implementa uma topologia retilínea 3-D. Permite obter respostas impulsivas (RIR) recorrendo a um processo iterativo que implementa as equações que descrevem o comportamento dos nós numa malha de guias de onda digitais. Como se pode observar na figura 3.5, os dados de entrada do *DWsim* são um ficheiro de extensão *.sim*, outro de extensão *.mtl* e o número de iterações a realizar. Os resultados da simulação são apresentados sob a forma de um ficheiro de extensão *.res*. Os ficheiros envolvidos são explicados de forma mais detalhada nas secções seguintes.

#### 3.3.2 Configuração da Simulação

O ficheiro *.sim* contém a informação necessária ao processo de modelação. Trata-se de um ficheiro em formato binário composto pelos seguintes campos:

- **tipo** *unsigned char* - tipo de simulação (mono-aural ou bi-aural)
- **fa** *unsigned float* - frequência de amostragem
- **pfonte** *unsigned int* [3] - posição  $(x, y, z)$  da fonte
- **pesq** *unsigned int* [3] - posição  $(x, y, z)$  do ouvido esquerdo
- **pdir** *unsigned int* [3] - posição  $(x, y, z)$  do ouvido direito
- **m3dim** *unsigned int* [3] - dimensão da matriz 3D de pontos  $(Nx, Ny, Nz)$



- **m3d** *unsigned char* [  $Nx$  ][  $Ny$  ][  $Nz$  ] - matriz 3D que representa o espaço físico a modelar

Este último campo é o mais importante, pois define a configuração geométrica da sala. Cada elemento da matriz **m3d** representa o material encontrado no ponto correspondente da sala. Note-se que o tipo utilizado (*unsigned char*) permite representar até 256 tipos de materiais diferentes. No entanto, alguns caracteres são reservados, como veremos adiante. A leitura correcta da matriz **m3d** exige que se conheçam as suas dimensões, dadas no campo **m3dim**.

O campo **tipo** indica ao simulador se deve gerar uma RIR (mono-aural) ou uma BRIR (bi-aural). Sublinhe-se que, no caso da simulação bi-aural, o ouvinte é parte integrante do espaço físico a modelar. A frequência de amostragem, embora não relevante no processo de simulação, é um campo importante no sentido em que confere ao modelo uma dimensão física real (equação 3.7). A resposta gerada pelo modelo assume a frequência de amostragem especificada.

O modelo assume uma fonte pontual, cuja posição pode ser directamente indicada na matriz **m3d** na posição correspondente, pelo carácter reservado 'X'. Se essa indicação não existir, é assumida a posição indicada no campo do ficheiro **pfonte**.

A posição dos pontos de captação da resposta (ouvidos/microfone omni-direcional) pode ser também indicada na matriz de pontos. Os caracteres reservados para o efeito são o 'L' e 'R'. Na simulação bi-aural, o 'R' representa a posição do ouvido direito e 'L' representa a posição do ouvido esquerdo. Se estes caracteres não forem encontrados na matriz, o simulador assume os campos do ficheiro correspondentes às posições do ouvido esquerdo (**pesq**) e direito (**pdir**). No caso da simulação monoaural, o ponto de captação considerado é o indicado pelo carácter 'L' ou pelo campo **pesq**.

### 3.3.3 Caraterização Acústica dos Materiais

O ficheiro *.mtl* é responsável por atribuir um coeficiente de reflexão a cada material representado na matriz. Trata-se de uma tabela onde o utilizador pode inserir novos materiais e o respetivo coeficiente de reflexão. Esta abordagem facilita a atribuição de coeficientes de reflexão aos materiais utilizados sem que seja necessário modificar a matriz que representa o espaço.

### 3.3.4 Resultados da Simulação

Os resultados da simulação são fornecidos pelo *DWsim* sob a forma de um ficheiro com extensão *.res*. Este contém os seguintes campos de informação:

- **tipo** *unsigned char* - tipo de simulação (monoaural ou bi-aural)
- **fa** *unsigned float* - frequência de amostragem
- **niter** *unsigned int* - número de iterações realizadas
- **pfonte** *unsigned int* [3]- posição da fonte ( $xn, yn, zn$ )
- **pesq** *unsigned int* [3] - posição do ouvido esquerdo ( $xn, yn, zn$ )
- **pdir** *unsigned int* [3] - posição do ouvido direito ( $xn, yn, zn$ )

- **rir/brir** *float* [niter/2·niter] - resposta impulsional monoaural/bi-aural

O campo **rir/brir** contém o resultado da simulação. No caso da simulação mono-aural, é uma RIR com o número de elementos igual ao número de iterações realizadas (**niter**). No caso de simulação bi-aural, será uma BRIR, com o dobro dos elementos. Os restantes campos não resultam de qualquer cálculo, apenas reproduzem informação de entrada sobre a posição dos ouvidos (ou microfone omni-direcional) e da fonte na simulação efetuada. Em simulações mono-aurais, o campo **pdir** é desprezado.

### 3.3.5 Algoritmo de Simulação

O modelo consiste numa malha retilínea de nós. Cada nó representa o ponto da sala correspondente e coneta-se aos seus seis nós vizinhos por ligações bidirecionais; tem, portanto, seis portos de saída e seis portos de entrada. A figura 3.6 representa um nó pertencente à malha. Assim, pode ser representado por uma estrutura de dados com a seguinte forma:

```
struct node{
    float p;                                /* valor de pressão no nó */
    /* portos de saída */
    float up_out;                           /* porto de cima */
    float down_out;                         /* porto de baixo */
    float front_out;                       /* porto da frente */
    float back_out;                        /* porto de trás */
    float right_out;                       /* porto da direita */
    float left_out;                        /* porto da esquerda */
    /* portos de entrada */
    float up_in;                           /* porto de cima */
    float down_in;                         /* porto de baixo */
    float front_in;                       /* porto da frente */
    float back_in;                        /* porto de trás */
    float right_in;                       /* porto da direita */
    float left_in;                        /* porto da esquerda */
    /* tipo de material */
    char material;
};
```

A variável *p* contém o valor da pressão do nó a cada iteração. O tipo de material que o nó representa é definido pela variável *material*. As restantes variáveis constituem os portos de comunicação. A malha de nós pode então ser definida como uma matriz tridimensional de elementos do tipo *node* (nó).

Numa primeira fase, o *DWsim* reserva memória para o modelo (malha de nós) e inicializa a variável *material* em todos os nós. Essa informação é obtida a partir da matriz *m3d* contida no ficheiro *.sim*. Se nenhum problema for encontrado (e.g. não indicação da posição da fonte e dos pontos de captação, dimensões inconsistentes com o número de elementos da matriz...) a simulação é iniciada.

Cada iteração do algoritmo compreende dois passos distintos. O primeiro - *scattering* - aplica a equação 3.6 para calcular, a partir dos valores nos portos de entrada, o valor de pressão em cada nó e em cada porto de saída. O segundo - *delay* - coloca o valor dos portos de saída de cada nó, nos portos de entrada dos nós vizinhos correspondentes. No passo *scattering*, os nós-fronteira são tratados de forma especial; o valor em cada porto de saída é calculado multiplicando o valor do porto de entrada correspondente pelo coeficiente de reflexão do material.

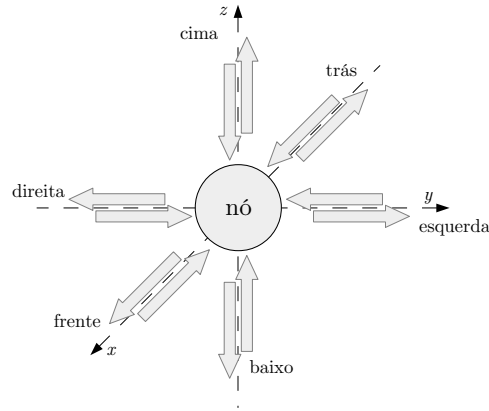


Figura 3.6: Representação tridimensional de um nó da malha.

Apenas os nós-fronteira (que estão conectados a nós de ar) e os nós de ar têm influência na modelação. Os restantes nós podem ser considerados inativos, pois não influenciam a simulação: são representados pelo carácter reservado 'I'. Isto significa que são ignorados durante o cálculo, melhorando assim a eficiência do simulador. O processo iterativo da simulação é descrito em termos gerais no seguinte código:

```
for(int n=0;n<niter;n++){
    /* insere amostra do sinal de entrada */
    put_input();

    /* calcula as saídas de cada nó */
    scattering_pass();

    /* copia as saídas de cada nó para as entradas dos nós vizinhos */
    delay_pass();

    /* regista o valor dos nós de medição */
    get_output();
}
```

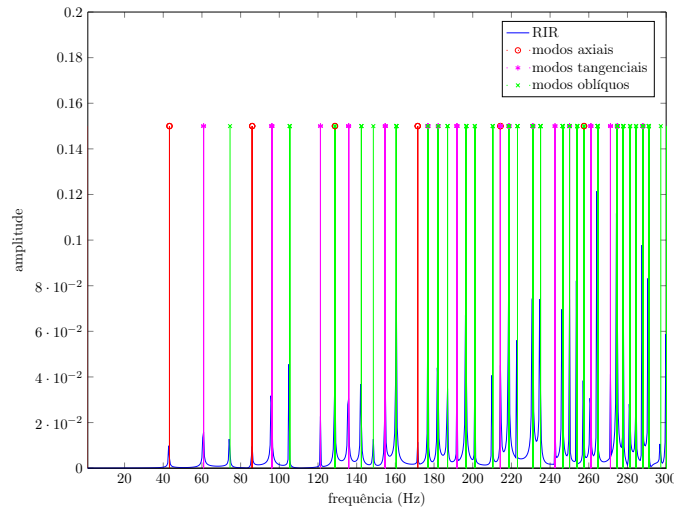
O código fonte completo do *DWsim* pode ser consultado no anexo B.

### 3.3.6 Validação

Os modos da resposta em frequência de um volume regular de dimensões  $L \times W \times H$  (forma paralelepípedica) podem ser calculados analiticamente pela seguinte expressão:

$$f_r = \frac{c}{2} \cdot \sqrt{\left(\frac{p}{L}\right)^2 + \left(\frac{q}{W}\right)^2 + \left(\frac{r}{H}\right)^2} \quad (3.8)$$

onde  $f_r$  representa a frequência de ressonância e  $p$ ,  $q$  e  $r$ , números inteiros maiores ou iguais a zero. Podem ser distinguidos três tipos de modos de ressonância: axial, tangencial e oblíquo. O



*Figura 3.7: Modos de ressonância teóricos e simulados.*

objetivo é comparar os modos teóricos calculados com picos da RIR no domínio da frequência. Para isso, foi simulada uma sala cúbica com 4 metros de lado. A fonte foi colocada num canto da sala e o microfone omnidirecional (simulação mono-aural) colocado no canto oposto. Foi utilizada uma frequência de amostragem de 44100Hz e realizadas 30000 iterações.

A figura 3.7 apresenta os modos calculados teoricamente através de uma risca na frequência correspondente a cada modo. É também apresentada na mesma figura, a RIR obtida por simulação para efeitos de comparação.

Note-se que, de facto, os picos da RIR em frequência coincidem com os modos calculados. Isto indica que o simulador implementa corretamente o algoritmo de modelação por guias de onda digitais.

## Capítulo 4

# Auralização em Tempo Real

### 4.1 Introdução

Auralização é o processo que permite a um utilizador escutar os estímulos auditivos produzidos por fonte(s) sonora(s) num espaço virtual (neste caso, uma sala). Esses estímulos correspondem à convolução entre o(s) sinal(is) emitido(s) pela(s) fonte(s) (gerado(s) em condições anecóicas) e a resposta impulsional bi-aural do espaço (BRIR). Normalmente, esta é obtida através de modelos acústicos, pelo que o processo de auralização é uma sequência de duas operações:

1. obtenção da BRIR por modelação;
2. convolução entre BRIR e sinal de entrada;

A auralização em tempo real consiste em realizar estas duas operações e assim gerar e escutar os estímulos de saída à mesma cadência a que o sinal de entrada é recebido. Devido à elevada carga computacional necessária à obtenção do modelo, isto torna-se impraticável. A solução passa por modelar previamente o espaço em causa, obtendo-se assim as respostas impulsionais bi-aurais (BRIR) para as posições que se pretende auralizar. Assim, a auralização em tempo real passa a depender apenas do tempo necessário à operação de convolução entre o sinal anecóico e a resposta impulsional bi-aural para uma determinada posição do ouvinte e da fonte. Graças à enorme evolução tecnológica nas recentes décadas e à utilização de algoritmos sofisticados, é hoje perfeitamente possível realizar a convolução em tempo real, mesmo com RIR muito longas, em computadores de uso geral.

No capítulo anterior foi apresentado e explicado o funcionamento do *DWsim*. Neste capítulo, pretende-se modelar previamente uma sala por forma a obter as BRIR correspondentes a um trajeto pré-definido do ouvinte. As BRIR obtidas são utilizadas para auralizar em tempo real a sala nesse mesmo trajeto recorrendo ao método de convolução *Overlap-Add* explicado na secção seguinte.

### 4.2 Convolução pelo Método *Overlap-Add*

A convolução é uma operação pesada computacionalmente. No entanto, é possível realizá-la em tempo real. Um dos algoritmos mais utilizados para este efeito é o designado *OverlapAdd*

- *OLA*. Este permite que o sinal de auralização seja obtido por blocos, sendo apenas realizadas as operações estritamente necessárias para gerar o bloco de saída corrente. Além disto, oferece a vantagem adicional de acomodar de forma eficiente transições entre respostas impulsiais. Repare-se que a operação de convolução com todas as amostras da resposta impulsional geraria um sinal de saída cujas amostras mais tardias não seriam necessárias ao processo de auralização caso ocorresse transição entre RIR, o que se traduziria em desperdício de recursos.

O método *OLA* é descrito em [4] e a sua explicação é aqui apresentada sucintamente. Como nota prévia, diga-se que o processo de auralização exige convolucionar o sinal de entrada (áudio) com cada uma das RIR que compõem a BRIR (uma para cada ouvido). Por uma questão de simplicidade, nesta explicação, apenas se considera a RIR correspondente a um dos ouvidos (o esquerdo). A operação é análoga para o outro ouvido.

Considere-se então, para uma determinada posição do ouvinte e fonte, a resposta impulsional (RIR) relativa ao ouvido esquerdo,  $h_L(n)$ , sendo  $x(n)$  o sinal de áudio emitido pela fonte virtual. O sinal de saída,  $y_L(n)$ , para auralização nessa posição é dado por:

$$y_L(n) = x(n) * h_L(n) \quad (4.1)$$

A soma de convolução (símbolo  $*$ ) é definida por:

$$y_L(n) = x(n) * h_L(n) = \sum_{k=-\infty}^{+\infty} x(k) \cdot h_L(n - k) \quad (4.2)$$

A convolução é, de facto, uma operação muito pesada, sobretudo se a RIR tiver um elevado número de amostras. Pode-se beneficiar do facto de esta poder ser realizada no domínio da frequência: sejam  $X(f)$  e  $H_L(f)$  as transformadas discretas de *Fourier* (DFT) para os sinais  $x(n)$  e  $h_L(n)$ , respetivamente:

$$x(n) \xrightarrow{DFT} X(f) \quad (4.3)$$

$$h_L(n) \xrightarrow{DFT} H_L(f) \quad (4.4)$$

Então, a transformada em frequência do sinal de saída,  $y_L(n)$ , é dada por:

$$Y_L(f) = X(f) \cdot H_L(f) \quad (4.5)$$

Naturalmente, a operação de multiplicação exige um tempo de computação muito inferior ao necessário para realizar a convolução no domínio do tempo. No entanto, é necessário que os sinais a convolucionar se encontrem no domínio da frequência. Uma vez que as respostas impulsiais são conhecidas à partida, as suas transformadas podem ser obtidas previamente, poupando-se assim, tempo de computação. Assim, a convolução entre  $x(n)$  e  $h_L(n)$  pode ser obtida por uma operação de multiplicação seguida do cálculo da transformada inversa (iDFT)

do sinal  $Y_L(f)$ :

$$Y_L(f) \xrightarrow{iDFT} y_L(n) \quad (4.6)$$

A operação de convolução realizada no domínio da frequência e o pré-cálculo das transformadas das RIR tornam o método *OLA* muito eficiente computacionalmente.

A figura 4.1 ilustra o método *OLA* na obtenção de um bloco de saída. A RIR é dividida em blocos de  $N$  amostras e o sinal de entrada,  $x(n)$ , é também processado em blocos com o mesmo número de amostras. O  $m$ -ésimo bloco do sinal de entrada,  $x(n)$ , é obtido por:

$$x_{Bm} = x[mN : (m+1)N - 1] \quad (4.7)$$

e o  $m$ -ésimo bloco da RIR,  $h(n)$ , por:

$$h_{Bm} = h[mN : (m+1)N - 1] \quad (4.8)$$

O número de blocos que constituem uma RIR com  $L$  amostras é definido por:

$$NB = \frac{L}{N} \quad (4.9)$$

O número de blocos  $NB$  tem, naturalmente, que ser um número inteiro. Para isso, é necessário que o comprimento  $L$  da RIR seja múltiplo inteiro de  $N$ . Este ajuste pode ser conseguido adicionando zeros à RIR (*zero-padding*) ou escolhendo um valor de  $N$  que satisfaça esta condição.

O bloco de áudio mais recente,  $x_{Bm}$ , é convolucionado com o primeiro bloco da RIR,  $h_0$ ; o bloco de áudio anterior,  $x_{B(m-1)}$ , é convolucionado com o segundo bloco da RIR,  $h_1$ , e assim sucessivamente para todos os  $NB$  blocos da RIR. A soma das convoluções origina um bloco intermédio com  $2N - 1$  amostras:

$$y'_{Bn} = \sum_{l=0}^{NB-1} x_{B(l)} * h_{B(n-l)} \quad (4.10)$$

Note-se que a convolução entre dois sinais com  $N$  amostras resulta num sinal com  $2N - 1$  amostras. O bloco intermédio gerado correspondente ao bloco de áudio mais recente é somado ao bloco intermédio obtido por operação análoga na iteração anterior da seguinte forma:

$$y_{Bn} = \begin{cases} y'_{B(n-1)}[n+N] + y'_{Bn}[n] & 0 \leq n < N-1 \\ y'_{Bn}[n] & n = N-1 \end{cases} \quad (4.11)$$

O bloco de saída corrente,  $y_{Bn}$ , formado por  $N$  amostras, combina o resultado da convolução entre o último bloco de áudio recebido,  $x_{Bm}$ , e a RIR, com a cauda ( $N - 1$  amostras) da convolução entre o penúltimo bloco de áudio e a RIR. Nesta lógica, as restantes amostras,

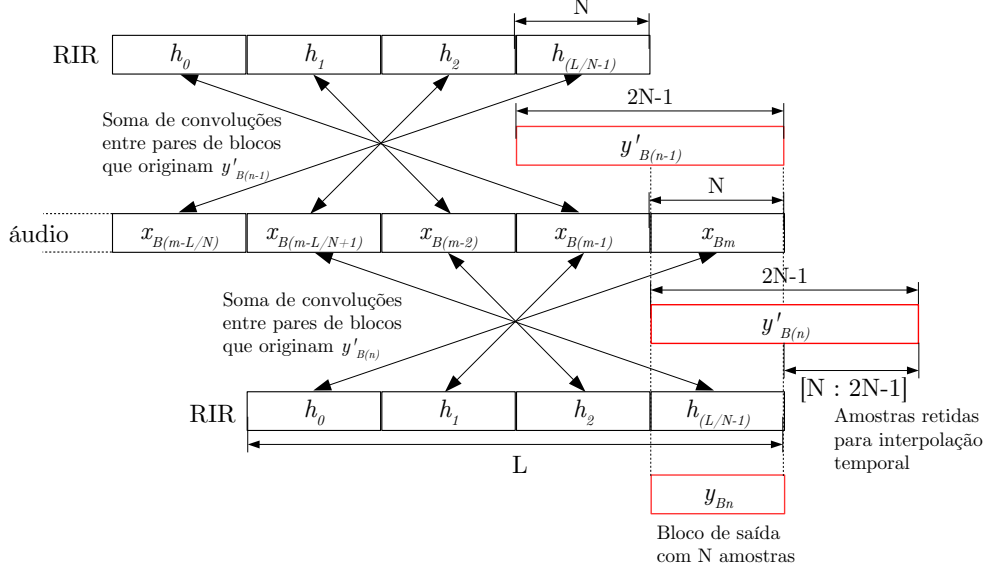


Figura 4.1: Esquema de funcionamento do método Overlap-Add [4]

$y'_{Bn}[N : 2N - 1]$ , são guardadas para produzir o próximo bloco de saída.

São necessários *buffers* de  $NB \cdot N$  amostras para guardar os blocos do sinal de entrada e da RIR. Note-se que para gerar o bloco de saída corrente, apenas é necessário conhecer a RIR e os últimos  $NB$  blocos do sinal de entrada. Isto significa que o sinal de entrada pode ser processado à medida que é recebido, gerando-se o bloco de saída correspondente.

O método *OLA* pode ser implementado no domínio da frequência. Isto significa que as operações de convolução são substituídas por operações de multiplicação, favorecendo assim a eficiência do método. As somas que produzem os blocos intermédios são efetuadas ainda no domínio da frequência. Naturalmente, é necessário, após isso, obter a transformada inversa do bloco para o domínio do tempo. É também necessário calcular a transformada em frequência de cada bloco recém-chegado do sinal de entrada. Isto implica um *over-head* de processamento adicional. No entanto, ele é compensado pela elevada eficiência da operação de multiplicação face à operação de convolução no domínio do tempo. Sublinhe-se que, como já foi referido, a transformada dos blocos da RIR pode ser realizada previamente.

As operações no domínio da frequência exigem que as DFT dos blocos do sinal de entrada e da RIR sejam realizadas com  $2N - 1$  amostras. Para isso, cada bloco de  $N$  amostras do sinal de entrada e da RIR é preenchido com zeros (*zero-padding*) até perfazer  $2N - 1$  amostras antes de ser calculada a sua transformada de *Fourier* discreta. Isto é necessário para que os blocos intermédios sejam constituídos por  $2N - 1$  amostras. Em cada interação, o bloco intermédio calculado é convertido para o domínio do tempo e somado ao bloco anterior, como



já foi explicado a propósito da equação 4.11.

### 4.3 Transição entre RIR e Interpolação Temporal

Considere-se agora a situação em que as respostas impulsionais, RIR, são alteradas durante o processo *OLA*, refletindo uma alteração da posição do ouvinte. A mudança entre RIR pode implicar um artefacto do ponto de vista da auralização, pois o bloco de saída correspondente à posição atual e o bloco correspondente à posição seguinte podem apresentar características bem distintas do ponto de vista auditivo. Este problema é também abordado em [4], onde é sugerido que se aplique interpolação temporal (*crossfade*) entre blocos de saída consecutivos. A ideia é obter um bloco de saída correspondente ao que se ouviria mantendo a RIR atual (posição atual) - bloco *fade-out* - e um bloco de saída correspondente ao que se ouviria com a RIR seguinte (posição seguinte) - bloco *fade-in*. Isto implica o triplo das operações. Além do cálculo do bloco intermédio para a posição atual (que é somado ao bloco intermédio anterior para formar o bloco de saída atual), é necessário calcular os blocos intermédios relativos à posição seguinte: um bloco intermédio relativo às  $N$  amostras seguintes do sinal de entrada e outro bloco relativo às  $N$  amostras atuais. É portanto, necessário guardar os últimos  $NB + 1$  blocos do sinal de entrada para calcular o bloco de *fade-in* e de *fade-out*.

Repare-se que quando não há transição entre RIR, os blocos de saída (*fade-out* e *fade-in*) são iguais e o resultado da interpolação temporal é correcto. Assim, a interpolação pode ser realizada sempre, independentemente de haver ou não mudança de RIR. O processamento adicional exigido pela implementação da interpolação temporal não põe em causa a auralização em tempo real.

### 4.4 Implementação Prática

Foi desenvolvida, em *Matlab*, uma biblioteca, denominada *OLAlib*, constituída por funções que permitem implementar o método *OLA* com interpolação temporal. Antes de serem apresentadas as funções desenvolvidas, é necessário conhecer a estrutura de dados que contém toda a informação sobre o estado do processo. Esta estrutura de dados, denominada *OlaStrct*, é atualizada a cada iteração e contém a informação necessária para que seja gerado o bloco de saída. A figura 4.2 apresenta os campos que a constituem.

As funções que compõem a biblioteca *OLAlib* são as seguintes:

- **ola\_initstrct** - inicializa a estrutura de dados *OlaStrct*. Tem como parâmetros de entrada o número de amostras pretendidas por bloco,  $N$ , e a RIR (em frequência) correspondente à posição inicial do ouvinte. Devolve a estrutura *OlaStrct* inicializada.
- **ola\_input2fft** - converte amostras do sinal de entrada para o domínio da frequência. Tem como parâmetros de entrada a estrutura *OlaStruct* e o sinal de entrada. Devolve um *buffer* com blocos de  $2N - 1$  amostras cada.
- **ola\_rir2fft** - converte uma RIR para o domínio da frequência. Tem como parâmetros de entrada o número de amostras por bloco  $N$ , e a RIR a converter. Devolve um *buffer* com blocos de  $2N - 1$  amostras cada.

```

% buffer circular onde são guardados os NB+1 blocos do sinal de entrada
% (em frequência)
OlaStrct.Binput

% buffer onde são guardados os NB blocos da RIR actual
% (em frequência)
OlaStrct.Brir_p

% buffer onde são guardadas as 2N-1 amostras do bloco intermédio actual
OlaStrct.y_pp

% bloco de fade-in resultante da RIR seguinte
% (no tempo)
OlaStrct.CrossBlk

% bloco actual resultante da RIR actual
% (no tempo)
OlaStrct.OutputBlk

% número de amostras por bloco
OlaStrct.N

% número de blocos que constituem a RIR
OlaStrct.NB

```

Figura 4.2: Estrutura de dados *OlaStrct*

- 
- **ola\_getinput** - adiciona um bloco (obtido pela função **ola\_input2fft**) ao *buffer OlaStrct.Binput* para ser processado na próxima iteração. Tem como parâmetros de entrada a estrutura *OlaStrct* e o bloco a ser adicionado. Devolve a estrutura *OlaStrct* atualizada.
  - **ola\_process** - calcula os blocos de saída (*fade-in* e *fade-out*) e devolve o resultado no campo *OutputBlk* da estrutura de dados *OlaStrct*. Tem como parâmetros de entrada a estrutura *OlaStrct* e o *buffer* (obtido pela função **ola\_rir2fft**) relativo da posição seguinte.

A utilização da biblioteca *OLAlib* facilita a implementação do método *OLA*, como evidencia o exemplo apresentado na figura 4.3.

Para validação, compararam-se resultados de convolução utilizando o método *OLA* implementado pela biblioteca *OLAlib* com os resultados obtidos para sinais idênticos com funções de convolução *standard* do *Matlab*. Concretamente, foi utilizado um conjunto de sinais gerados aleatoriamente pelo *Matlab* com diferentes números de amostras. É importante salientar que foram utilizados diferentes números de amostras por bloco,  $N$ , nas diversas comparações. Verificou-se que os resultados obtidos pelos métodos foram sistematicamente iguais, indicando assim que a biblioteca desenvolvida implementa o método *OLA* correctamente. O código-fonte da biblioteca *OLAlib* pode ser consultado no anexo C.

```

% define o número de amostras por bloco
N = 1024;

% converte a resposta impulsional  $h(n)$  para o domínio da frequência
%  $h(n) \rightarrow$  vector  $h$ 
hFT = ola_rir2fft(N, h);

% inicializa a estrutura de dados
OlaStrct = ola_initstrct(N, hFT);

% converte o sinal de entrada  $x(n)$  para frequência
%  $x(n) \rightarrow$  vector  $x$ 
xFT = ola_input2fft(OlaStrct, x);

% processa até  $x(n)$  terminar
while(x!= end)
{
    % adiciona bloco do sinal de entrada
    OlaStrct = ola_getinput(OlaStrct, xFT);

    % calcula bloco de saída correspondente
    OlaStrct = ola_process(OlaStrct, hFT);

    % bloco de saída
    output_block = OlaStrct.OutputBlk;
}

```

Figura 4.3: Pseudo-código da implementação do método OLA utilizando a biblioteca OLALib.

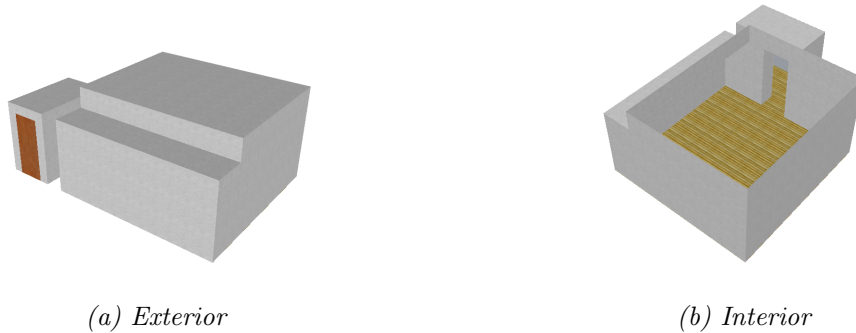


Figura 4.4: Representação tridimensional da sala utilizada no estudo de caso

## 4.5 Auralização num Trajeto Virtual

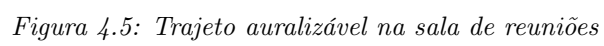
Nesta secção é estudado um caso particular de auralização de uma sala com um trajeto pré-definido do ouvinte. O *DWsim* permite obter o modelo da sala para uma determinada posição do ouvinte e da fonte. A biblioteca *OLALib* proporciona a auralização da sala, em tempo real, a partir das respostas impulsivas bi-aurais geradas pelo *DWsim*. Foi escolhida uma sala de dimensões relativamente reduzidas. Trata-se de uma sala de reuniões do Instituto de Engenharia Electrónica e Telemática de Aveiro (IEETA) cuja representação 3-D se encontra ilustrada na figura 4.4. A matriz de pontos da sala foi obtida por métodos automáticos de aquisição de dados 3-D no âmbito de uma outra dissertação [5]. A distância internodal utilizada corresponde a uma frequência de amostragem de  $44100Hz$ . A figura 4.5 apresenta o esquema de auralização idealizado para o ouvinte.

O esquema de auralização idealizado exigiu a obtenção de dezoito respostas impulsivas bi-aurais (BRIR) - uma por cada posição do ouvinte. A distância entre consecutivas BRIR é de aproximadamente  $50cm$ . A posição da fonte sonora é fixa. Para cada um dos processos de modelação foi gerado um ficheiro *.sim* cuja matriz de pontos representa a sala, a fonte e o ouvinte nas posições pretendidas.

A representação adotada para a cabeça do ouvinte consiste numa matriz cúbica de pontos, sendo os que fazem parte de uma superfície esférica de raio,  $r_n$ , assinalados como pontos representativos do material da cabeça - pele. A figura 4.6 ilustra a representação em causa. A distância internodal da matriz de pontos da cabeça é, naturalmente, a mesma utilizada na matriz de pontos da sala. Os ouvidos encontram-se equidistantes do centro da esfera e a sua posição define a orientação da cabeça.

Foi criada uma base de dados de cabeças, com cerca de  $13cm$  de diâmetro, orientadas em todas as posições possíveis. Para isso, foi utilizada a mesma representação da cabeça e apenas foram alteradas as posições dos ouvidos. Isto significa que apenas valores discretos de azimute são permitidos. Esta abordagem facilita a inclusão do ouvinte na sala, embora a representação seja aproximada. A matriz tridimensional da cabeça é inserida ponto-a-ponto na matriz tridimensional da sala, na posição e orientação pretendidas para o ouvinte. Este processo foi repetido para as dezoito posições.

A representação da sala considera apenas dois tipos de materiais: madeira (chão) e paredes (betão pintado). Estes materiais são representados na matriz de pontos da sala pelos caracteres 'M' e 'P' respetivamente. O material da cabeça - pele - é representado pelo carácter 'H'. Os pontos referentes ao ouvido esquerdo e direito são assinalados com o carácter reservado 'L' e



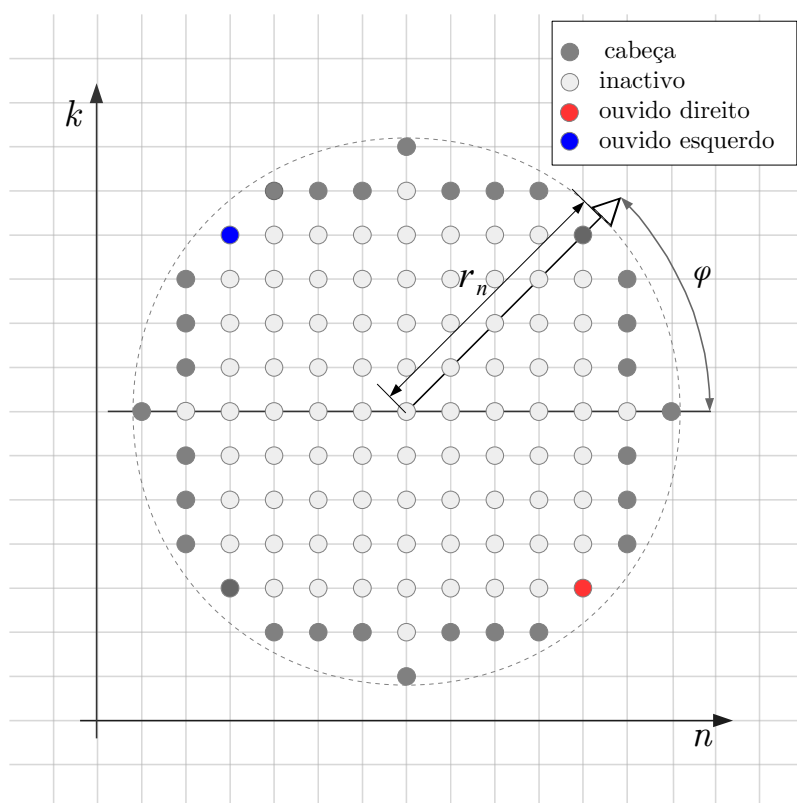


Figura 4.6: Representação da cabeça do ouvinte por uma matriz tridimensional de pontos.

'R' respetivamente. Foram então gerados os ficheiros *.sim* correspondentes a cada posição do ouvinte na sala. A posição da fonte foi assinalada no campo correspondente do cabeçalho do ficheiro (não foi diretamente assinalada na matriz com o carácter 'X'). O número de iterações realizado por cada simulação foi 44100 (BRIR com 1 segundo).

Cada simulação demorou cerca de 17 horas e ocupou cerca de 2GB de memória RAM. A máquina utilizada no processo de simulação tem um processador *Intel(R) Core(TM) i7-3770 CPU* operando a 3.40 GHz.

Uma vez obtidos os ficheiros de saída, resultantes das simulações realizadas, foi criada uma aplicação *Matlab* que, recorrendo à biblioteca *OLAlib*, permite auralizar o trajeto do ouvinte na da sala. Neste aplicação, o utilizador pode usar o teclado do computador para transitar entre posições do ouvinte. A figura 4.7 ilustra o aspecto da aplicação desenvolvida. Note-se que a BRIR só é conhecida em pontos discretos (18 neste caso); por isso, em cada ponto de percurso é utilizada a BRIR disponível que corresponder ao ponto mais próximo. Isto resulta em transições algo abruptas (embora sem *'clicking'*, graças à interpolação temporal). No entanto, os resultados são consistentes no sentido em que é percetual a posição da fonte ao longo do trajeto. No anexo D podem ser encontrados diversos ficheiros áudio correspondentes ao percurso do ouvinte na sala, para diferentes tipos de sinais de entrada.

A limitação de movimento do ouvinte durante o processo de auralização pode ser resolvida obtendo-se BRIR para posições adicionais. No entanto, isto tem um custo computacional elevadíssimo. A interpolação espacial entre BRIR é, intuitivamente, uma forma de contornar esta limitação. Utilizando as BRIR mais próximas, pode ser possível obter uma BRIR sintetizada para qualquer ponto do trajeto. Este tema é abordado no capítulo seguinte.

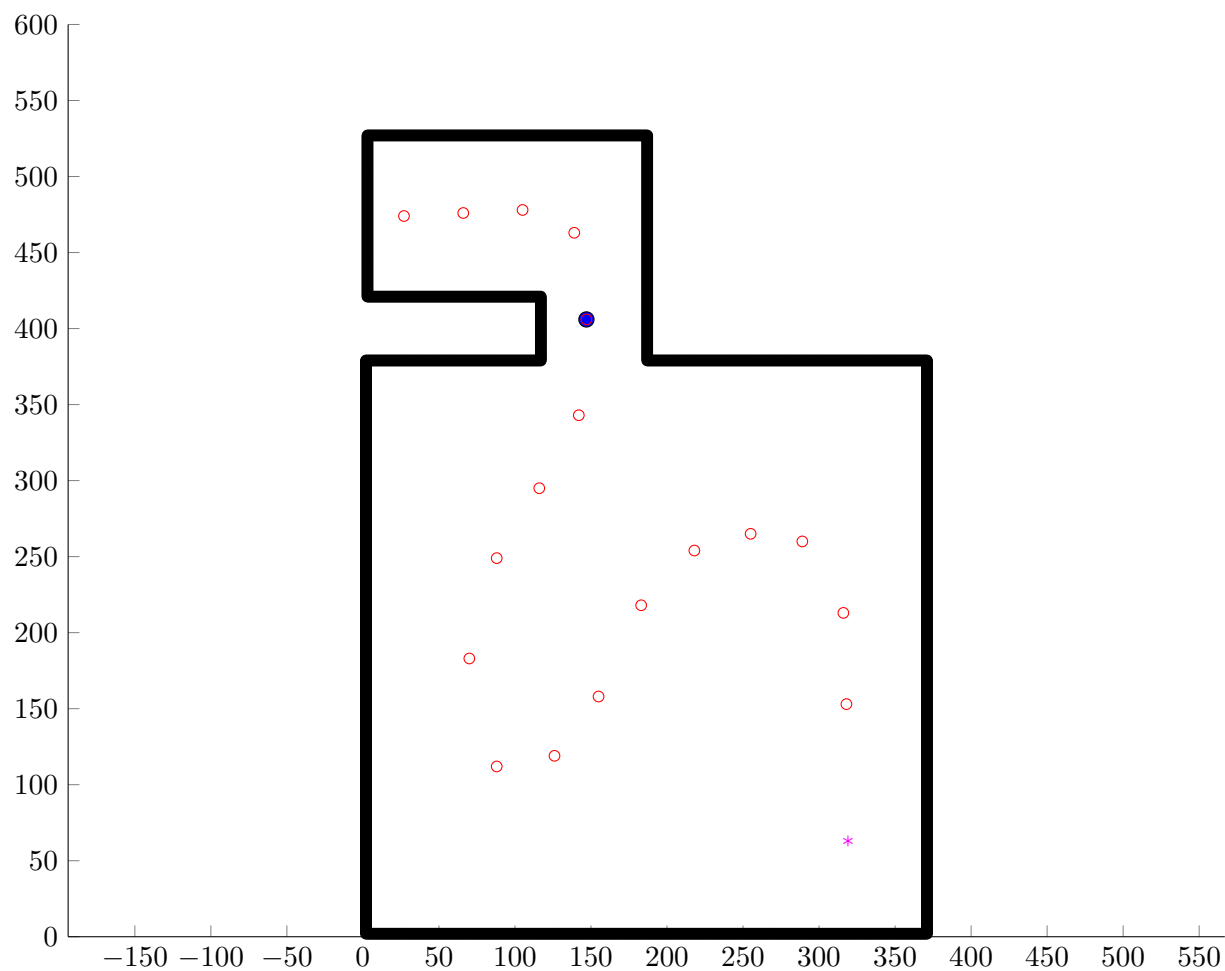


Figura 4.7: Aplicação (em Matlab) que permite auralizar em tempo real a sala considerada no estudo de caso.

Os círculos vermelhos representam as posições para as quais foram obtidas as BRIR. O ouvinte é representado por uma circunferência azul e a sua orientação é sempre tangencial ao percurso. A fonte é assinalada pelo asterisco rosa. As dimensões da sala são apresentadas em número de nós.



## Capítulo 5

# Interpolação de Respostas Impulsionais

### 5.1 Introdução

No capítulo anterior foi implementada a auralização em tempo real para o caso de estudo proposto. Foi definido o percurso do ouvinte no interior da sala e obtidas 18 respostas impulsionais bi-aurais (BRIR) correspondentes a pontos ao longo desse percurso espaçadas cerca de 50cm. Na auralização com estes dados, são perceptíveis mudanças repentinas causadas pelas transições entre BRIR. A obtenção de mais respostas impulsionais pode atenuar este problema. Contudo, o tempo de simulação aumentaria em proporção. Como referido no capítulo anterior, foram necessárias 17 horas de simulação para cada uma das 18 BRIR obtidas no estudo de caso. A síntese de respostas impulsionais por interpolação entre as BRIR mais próximas parece, intuitivamente, uma solução a ter em conta para manter os tempos de cálculo dentro de limites praticáveis. Nesta secção, são estudados e testados processos de interpolação de respostas impulsionais.

### 5.2 Interpolação Linear

A abordagem mais imediata para sintetizar uma RIR é interpolar linearmente as BRIR mais próximas. A interpolação entre duas BRIR consiste num par de interpolações entre as duas RIR correspondentes a cada ouvido. Assim, a explicação apresentada refere-se apenas à interpolação entre as RIR de um dos ouvidos, sendo a outra realizada de forma análoga.

A interpolação linear consiste em realizar uma média ponderada entre as respostas impulsionais conhecidas mais próximas, como ilustra a figura 5.1.

A resposta impulsional  $h_I$  é obtida por:

$$h_I = \frac{d_B}{d_A + d_B} \cdot h_A + \frac{d_A}{d_A + d_B} \cdot h_B \quad (5.1)$$

Pela equação 5.1, verifica-se que  $h_I = h_A$  quando  $d_A = 0$  e  $h_I = h_B$  quando  $d_B = 0$ . Isto serve as condições-limite da interpolação em que a resposta interpolada corresponde à mesma

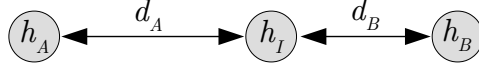


Figura 5.1: Esquema de interpolação linear de respostas impulsivas. A resposta no ponto intermédio, correspondente a  $h_I$ , resulta da média ponderada (em função de  $d_A$  e  $d_B$ ) das respostas conhecidas,  $h_A$  e  $h_B$ .

posição de uma das respostas originais.

Sendo a *Transformada de Fourier* uma operação linear, esta técnica de interpolação pode também ser aplicada no domínio da frequência, obtendo-se os mesmos resultados:

$$h_A \xrightarrow{\mathfrak{F}} H_A \quad (5.2)$$

$$h_B \xrightarrow{\mathfrak{F}} H_B \quad (5.3)$$

$$H_I = \frac{d_B}{d_A + d_B} \cdot H_A + \frac{d_A}{d_A + d_B} \cdot H_B \quad (5.4)$$

$$h_I \xleftarrow{\mathfrak{F}^{-1}} H_I \quad (5.5)$$

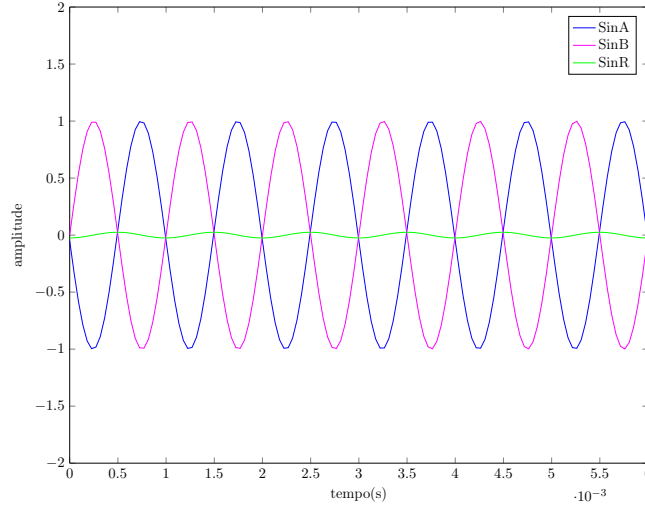
Para avaliar os efeitos da interpolação linear aqui apresentada, foi realizado um teste com duas sinusóides, segundo o esquema apresentado na figura 5.1. Considerou-se  $d_A = d_B$  (ponto médio). As sinusóides intervenientes na interpolação estão desfasadas e têm a mesma amplitude. A figura 5.2 apresenta os resultados obtidos. As sinusóides *sinA* e *sinB* originam por interpolação linear a sinusóide *sinR*. Note-se que a amplitude de *sinR* é muito reduzida comparada com as amplitudes de *sinA* e *sinB*. Isso acontece devido ao facto de o desfasamento entre sinusóides ser próximo de 180 graus (oposição de fase). Ocorre interferência destrutiva no processo de interpolação, o que resulta num sinal interpolado que não é o pretendido. Seria desejável que o resultado fosse uma sinusóide com a mesma amplitude de *sinA* e *sinB* e com um atraso de fase correspondente à média dos atrasos de fase de *sinA* e *sinB*. Por isso, a interpolação linear não serve o propósito pretendido.

Para analisar de uma perspetiva diferente esta mesma situação, consideremos o esquema apresentado na figura 5.3, onde é emitida uma onda (com uma determinada frequência  $f$ ) pela fonte sonora. O tempo que a onda demora a chegar às posições correspondentes a  $h_A$  e  $h_B$ , é representado por  $t_A$  e  $t_B$ , respetivamente. A diferença entre os atrasos de chegada pode ser definida por:

$$\Delta T = t_A - t_B \quad (5.6)$$

A diferença de atrasos,  $\Delta T$ , corresponde a uma diferença de distâncias percorridas dada por:

$$\Delta d = c \cdot \Delta T \quad (5.7)$$



*Figura 5.2: Interação linear entre duas sinusóides desfasadas com frequência e amplitude iguais. SinA e SinB representam as sinusóides interpoladas. SinR é a sinusóide resultante da interação linear.*

onde  $c$  representa a velocidade de propagação sonora no meio e  $\Delta d$  a diferença entre as distâncias percorridas. A velocidade de propagação, para um determinado meio, é definida por:

$$c = \lambda \cdot f \quad (5.8)$$

em que  $\lambda$  e  $f$  representam o comprimento de onda e a frequência da onda que se propaga no meio, respetivamente. Substituindo a equação 5.8 na equação 5.7, obtemos:

$$\frac{\Delta d}{\lambda} = f \cdot \Delta T \quad (5.9)$$

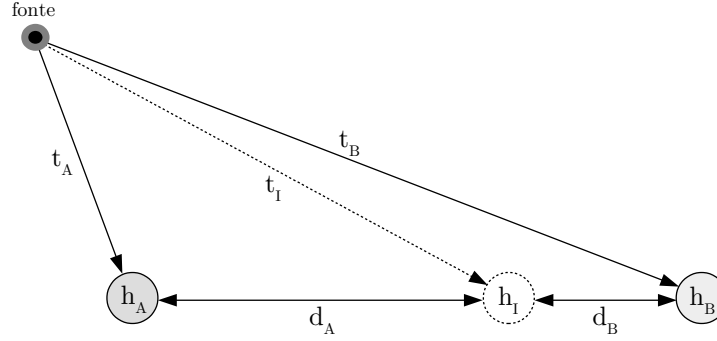
A interferência destrutiva ocorre devido à oposição de fase entre as ondas intervenientes na operação de interação linear. Sempre que:

$$\Delta d = \lambda n + \frac{\lambda}{2}, \quad n \in \mathbb{N} \quad (5.10)$$

as ondas que chegam a  $h_A$  e a  $h_B$  encontram-se em oposição de fase. A seguinte equação permite prever os valores de  $\Delta T$  (diferença entre tempos de chegada das ondas a  $h_A$  e  $h_B$ ) para os quais isso acontece:

$$\Delta T = \frac{2n + 1}{2f} \quad (5.11)$$

Isto significa que os atrasos das ondas não são devidamente interpolados. Por consequência,



*Figura 5.3: Esquema de interpolação de respostas impulsivas na presença da fonte sonora.  $t_A$  representa o tempo que uma onda emitida pela fonte demora a chegar à posição correspondente a  $h_A$ .  $t_B$  representa o tempo que uma onda emitida pela fonte demora a chegar à posição correspondente a  $h_B$ .  $t_I$  representa o tempo que uma onda demoraria a chegar à posição correspondente à resposta impulsional  $h_I$ , obtida por interpolação entre  $h_A$  e  $h_B$ .*

podemos concluir que o mesmo acontece para as várias componentes em frequência que compõem as respostas impulsivas. Ou seja, a interpolação linear não permite interpolar atrasos temporais entre respostas impulsivas. No anexo E podem ser encontrados resultados, sob a forma de gráficos, onde é perceptível o efeito da interferência destrutiva resultante da interpolação linear.

Na sequência do trabalho realizado, foram estudados dois artigos que abordam este assunto, [6] e [7]. Ambos visam, por métodos diferentes, a desejável interpolação de atrasos temporais. O primeiro propõe uma abordagem no domínio do tempo em que os instantes de ocorrência dos principais picos de cada RIR (correspondentes ao som directo e às reflexões dominantes da sala) são registados em vetores; é realizada uma interpolação linear entre esses vetores, resultando em instantes interpolados para cada pico. É aplicado então um processo de *time warping* (distorção da escala de tempo) das respostas originais de forma a reposicionar os seus picos nos instantes assim obtidos. Só após esse processo é que as respostas são interpoladas linearmente. O segundo propõe que a resposta impulsional sintetizada seja obtida interpolando linearmente o módulo e fase das respostas impulsivas conhecidas. Note-se que, como um atraso temporal corresponde, no domínio da frequência, a uma alteração de fase, a interpolação de fase constitui um meio de responder ao objetivo pretendido. Decidimos explorar este segundo método.

### 5.3 Interpolação em Fase e Módulo

É proposto em [7] que a interpolação de respostas impulsivas seja realizada recorrendo à interpolação linear em fase e módulo das respostas impulsivas conhecidas. Considere-se o esquema de interpolação apresentado na figura 5.3. A resposta impulsional  $h_I$  pode ser obtida executando os seguintes passos:

- Em primeiro lugar, são obtidas as respostas impulsionais,  $h_A$  e  $h_B$ , no domínio da frequência:

$$h_A \xrightarrow{\mathfrak{F}} H_A \quad (5.12)$$

$$h_B \xrightarrow{\mathfrak{F}} H_B \quad (5.13)$$

$\xrightarrow{\mathfrak{F}}$  representa a operação correspondente à *Transformada de Fourier*.

- De seguida são interpolados linearmente os seus módulos e fases:

$$|H_I| = \frac{d_B}{d_A + d_B} \cdot |H_A| + \frac{d_A}{d_A + d_B} \cdot |H_B| \quad (5.14)$$

$$\psi(H_I) = \frac{d_B}{d_A + d_B} \cdot \psi(H_A) + \frac{d_A}{d_A + d_B} \cdot \psi(H_B) \quad (5.15)$$

$| \cdot |$  e  $\psi$  representam o módulo e a fase, respetivamente.

- A resposta interpolada, no domínio da frequência, é então dada por:

$$H_I = |H_I| \cdot e^{j \cdot \psi(H_I)} \quad (5.16)$$

- Por fim, é calculada a transformada inversa de *Fourier* obtendo-se o resultado da interpolação,  $h_I$ , no domínio do tempo:

$$h_I \xleftarrow{\mathfrak{F}^{-1}} H_I \quad (5.17)$$

$\xleftarrow{\mathfrak{F}^{-1}}$  representa a operação correspondente à *Transformada de Fourier* inversa.

A interpolação de fases é a operação que marca a diferença na interpolação de respostas impulsionais. De fato, uma interpolação de atrasos entre duas ondas é uma interpolação de fases no domínio da frequência. Considere-se o esquema apresentado na figura 5.3 onde um impulso é gerado pela fonte sonora. Os tempos que o impulso demora a chegar às posições correspondentes a  $h_A$  e  $h_B$  são, respetivamente,  $t_A$  e  $t_B$ . Os impulsos que chegam a cada uma das posições podem ser vistos como impulsos atrasados e as suas transformadas em frequência são:

$$\delta(t - t_A) \xrightarrow{\mathfrak{F}} e^{-j\omega t_A} \quad (5.18)$$

$$\delta(t - t_B) \xrightarrow{\mathfrak{F}} e^{-j\omega t_B} \quad (5.19)$$

A interpolação linear dos atrasos  $t_A$  e  $t_B$ , representados na figura 5.3, depende da distância entre as posições onde são medidos os impulsos em relação à fonte. A interpolação destes atrasos

no domínio do tempo pode ser obtida em função das distâncias  $d_A$  e  $d_B$  pela seguinte equação:

$$t_I = \frac{d_B}{d_A + d_B} \cdot t_A + \frac{d_A}{d_A + d_B} \cdot t_B \quad (5.20)$$

onde  $t_I$  representa o atraso interpolado. A sua *Transformada de Fourier* é:

$$\delta(t - t_I) \xrightarrow{\mathfrak{F}} e^{-j\omega(\frac{d_B}{d_A + d_B} \cdot t_A + \frac{d_A}{d_A + d_B} \cdot t_B)} \quad (5.21)$$

Por outro lado, a interpolação das fases dos impulsos resulta em:

$$\phi_I = \omega(\frac{d_B}{d_A + d_B} \cdot t_A + \frac{d_A}{d_A + d_B} \cdot t_B) \quad (5.22)$$

onde  $\phi_I$  representa a fase obtida por interpolação. Isto significa que:

$$e^{-j\phi_I} = e^{-j\omega(\frac{d_B}{d_A + d_B} \cdot t_A + \frac{d_A}{d_A + d_B} \cdot t_B)} \quad (5.23)$$

Prova-se assim que, para ondas com a mesma frequência  $\omega$ , o impulso obtido por interpolação linear dos atrasos no domínio do tempo é igual ao impulso resultante da interpolação de fases. Naturalmente, as fases interpoladas têm de ser desenroladas (*unwrapped*). Para ilustrar a necessidade desta operação, considere-se a interpolação entre duas sinusóides,  $s_1$  e  $s_2$ , dadas por:

$$s_1 = \sin(\omega t) \quad (5.24)$$

$$s_2 = \sin(\omega t + 2\pi) \quad (5.25)$$

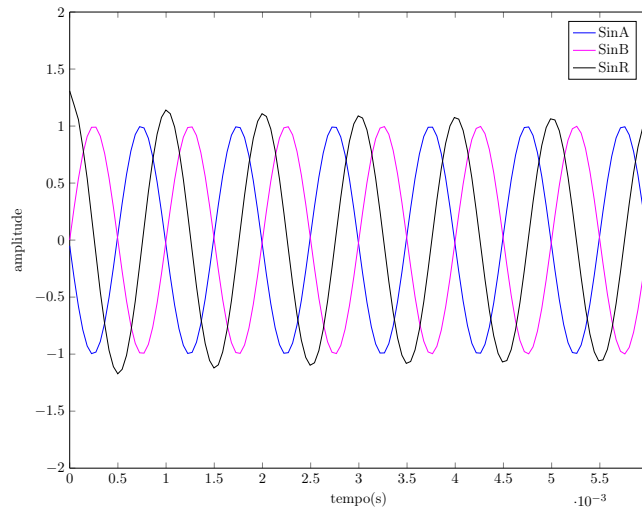
A sinusóide  $s_1$  está atrasada um período,  $2\pi$ , em relação a  $s_2$ . Seria de esperar que o resultado da interpolação de fases (no ponto intermédio) fosse uma sinusóide atrasada  $\pi$ . No entanto, por enrolamento de fases (*wrapping*), o argumento de  $s_1$  torna-se igual ao de  $s_2$  e o resultado será uma sinusóide em fase com  $s_1$  (e  $s_2$ ).

Frise-se que a interpolação linear em fase e módulo não corresponde a uma interpolação linear dos sinais interpolados. Isto é:

$$\alpha \cdot \rho_1 e^{-j\theta_1} + \beta \cdot \rho_2 e^{-j\theta_2} \neq (\alpha \cdot \rho_1 + \beta \cdot \rho_2) e^{-j(\alpha \cdot \theta_1 + \beta \cdot \theta_2)} \quad (5.26)$$

onde  $\alpha + \beta = 1$ .

Para se verificar a influência da interpolação em fase e módulo no domínio do tempo,



*Figura 5.4: Interpolação em módulo e fase entre duas sinusóides desfasadas com amplitude e frequência iguais.*

*SinA e SinB representam as sinusóides interpoladas. SinR é a sinusóide resultante da interpolação em módulo e fase.*

realizou-se um teste que consiste na interpolação de duas sinusóides *sinA* e *sinB* (analogamente ao teste realizado na secção anterior para a interpolação linear). Os resultados obtidos são apresentados nas figuras 5.4 e 5.5.

A figura 5.4 ilustra as sinusóides interpoladas *sinA* e *sinB* e a sinusóide resultante da interpolação em fase e módulo *sinR*. Observa-se que a sinusóide *sinR* tem uma amplitude aproximadamente igual à amplitude das sinusóides *sinA* e *sinB* e uma fase intermédia em relação às sinusóides interpoladas. Isto indica que este método de interpolação é capaz de interpolar os atrasos temporais entre dois sinais.

Foram realizados vários testes com sinusóides de diferentes amplitudes e fases (para se provar que o exemplo acima explicado não é apenas um caso particular da aplicação deste método de interpolação). Os resultados podem ser encontrados no anexo E.

## 5.4 Aplicação de Interpolação em Fase e Módulo

O método de interpolação em fase e módulo foi aplicado ao estudo de caso apresentado no capítulo anterior (auralização da sala num trajeto pré-definido). Foi implementada uma função que calcula as BRIR para cada posição do ouvinte na sala, efetuando interpolação, em módulo e fase, das BRIR conhecidas mais próximas (previamente obtidas por modelação física). O tempo de processamento extra necessário para efetuar estes cálculos não se revelou significativo; a auralização pôde continuar a ser realizada em tempo real.

Verificou-se, nos testes iniciais, que o som auralizado, nas posições das BRIR interpoladas, apresentava um eco com cerca de um segundo de atraso. Acabou por se verificar, experimentalmente, que a duração do eco estava directamente relacionada com a duração da BRIR. Ou seja, se forem adicionadas 44100 amostras nulas à BRIR o eco surge ao fim de dois segundos. Não há nenhuma informação sobre este fenómeno em [7]. Repare-se que se o número de amostras das BRIR sujeitas a interpolação for reduzido, o eco pode ser visto como reforço

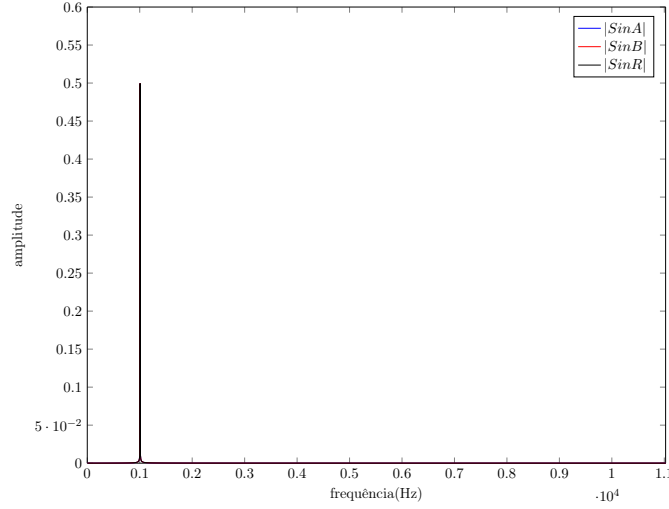


Figura 5.5: Módulo das sinusóides interpoladas e da sinusóide resultante da interpolação

do áudio auralizado. É de admitir que possa ser essa a causa de este fenómeno ter passado despercebido.

A figura 5.6 apresenta uma resposta impulsional sintetizada, resultante da interpolação em fase e módulo entre duas respostas impulsionais. A técnica de *unwrapping* de fases foi aplicada às respostas interpoladas. O sinal resultante da interpolação evidencia uma pequena elevação na cauda da resposta impulsional. Facilmente se percebe que a elevação é a razão da existência do eco. A análise da figura sugere imediatamente uma solução para este problema: cortar a parte da cauda correspondente à elevação prescindindo assim das reflexões mais tardias. Uma outra possível solução seria obter BRIR mais longas (por simulação) e cortar a cauda por forma a obter o número de amostras desejadas. Isto implicaria, naturalmente, um acréscimo do tempo de computação.

O resultado da interpolação entre duas RIR cortadas pode ser consultado no anexo F. Pode-se constatar que o eco se torna impercetível aplicando esta solução.

## 5.5 Uma Questão de Energia

O problema da interpolação de respostas impulsionais foi aparentemente resolvido ou pelo menos não é tão evidente do ponto de vista auditivo. No entanto, há um aspeto importante a ter em conta. Tomemos como exemplo a interpolação de duas respostas impulsionais segundo o esquema apresentado na figura 5.1. Seja  $Eh_I$  a energia do sinal resultante da interpolação,  $h_I$ , e  $Eh_A$  e  $Eh_B$  as energias dos sinais interpolados  $h_A$  e  $h_B$  respetivamente. É evidente que  $h_I$  terá uma contribuição energética de  $h_A$  e de  $h_B$  em função das distâncias  $d_A$  e  $d_B$ . Sendo  $\alpha = \frac{d_A}{d_A+d_B}$  e  $\beta = \frac{d_B}{d_A+d_B}$  a relação de energias deverá ser dada por:

$$E_{th} = \alpha \cdot Eh_A + \beta \cdot Eh_B \quad (5.27)$$

em que  $E_{th}$  é a energia teórica para a resposta interpolada  $h_I$ . No entanto isto não se verifica.



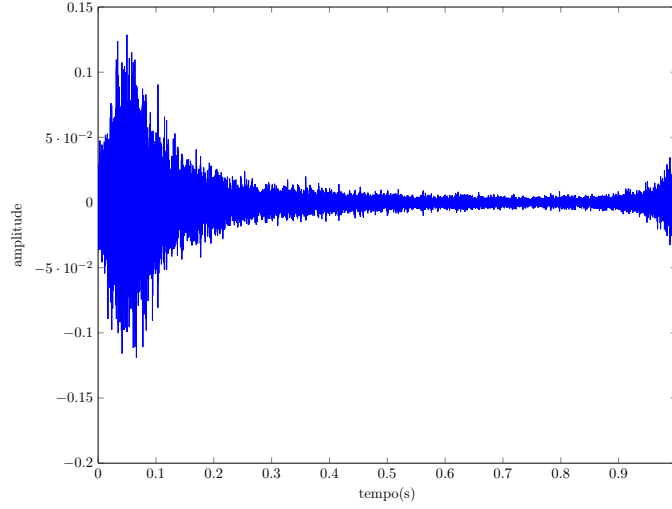


Figura 5.6: Resposta impulsional sintetizada por interpolação em fase e módulo com unwrap de fases

A energia do sinal interpolado não se relaciona pela equação 5.27. É possível obter uma relação entre a energia teórica e a energia de  $h_I$  obtido por interpolação. Essa relação de energias pode ser definida por:

$$\frac{E_{th}}{Eh_I} = K \quad (5.28)$$

em que  $K$  é constante e  $Eh_I$  a energia de  $h_I$ . A energia de um sinal discreto,  $x_n$ , é obtida por:

$$E_{x_n} = \sum_{n=1}^N x_n^2 \quad (5.29)$$

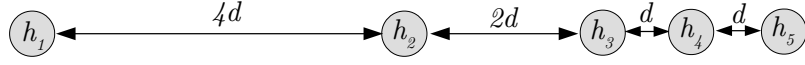
onde  $N$  é o número de amostras do sinal. Ou seja, a energia do sinal  $h_I$  é obtida por:

$$E_{h_I} = \sum_{n=1}^N h_I^2 \quad (5.30)$$

Substituindo 5.30 em 5.28 obtemos:

$$E_{th} = K \sum_{n=1}^N h_I^2 \Leftrightarrow E_{th} = \sum_{n=1}^N (\sqrt{K} \cdot h_I)^2 \quad (5.31)$$

Verifica-se, portanto, que o sinal resultante da interpolação deve ser multiplicado por um factor  $\sqrt{K}$  por forma a que as energias dos sinais intervenientes na interpolação sejam devidamente



*Figura 5.7: Esquema de teste para interpolação de respostas impulsivas.*

*As respostas impulsivas sintetizadas e simuladas são comparadas qualitativamente.  $h_2$  é comparada com a interpolação entre  $h_1$  e  $h_5$ .  $h_3$  é comparada com a interpolação entre  $h_2$  e  $h_5$ .  $h_4$  é comparada com a interpolação entre  $h_3$  e  $h_5$ .*

balanceadas. Isto é necessário na medida em que a energia da resposta impulsional obtida por interpolação tem que ser igual à soma das partes das energias das respostas interpoladas.

## 5.6 Testes

Uma vez que as técnicas apresentadas servem o propósito da interpolação de respostas impulsivas, foi montado um esquema de testes para tentar perceber até que ponto a distância entre respostas impulsivas conhecidas afeta a resposta impulsional sintetizada. Para isso, foram simuladas RIR (com 44100 amostras) numa sala cúbica com 2 metros de lado a uma frequência de  $44100\text{Hz}$ . A figura 5.7 ilustra o esquema de teste utilizado. O objetivo é comparar as RIR simuladas com as obtidas por interpolação (em fase e módulo). A RIR  $h_2$  pode ser comparada com a interpolação das RIR  $h_1$  e  $h_5$ . Da mesma forma,  $h_3$  pode ser comparada com a interpolação de  $h_2$  e  $h_5$ . Por fim,  $h_4$  pode ser comparada com  $h_3$  e  $h_5$ . No anexo G podem ser encontrados os resultados dos testes realizados.

A qualidade das respostas impulsivas sintetizadas pelos métodos apresentados ao longo do presente capítulo foi avaliada informalmente e por uma quantidade reduzida de sujeitos. No futuro, testes formais devem ser realizados para que se possam comparar devidamente respostas sintetizadas e simuladas.

## Capítulo 6

# Conclusões e Trabalho Futuro

### 6.1 Conclusões

No trabalho realizado, no âmbito desta dissertação, começou-se por estudar as técnicas de modelação acústica propostas na literatura. A modelação por guias-de-onda digitais (*DWM*) foi a técnica adoptada devido ao seu elevado rigor na descrição do comportamento sonoro em espaços físicos, nomeadamente salas. Neste sentido, foi desenvolvido um pacote de *software* que implementa modelação por guias-de-onda digitais - o *DWsim*. De seguida, foi desenvolvida uma biblioteca, em *Matlab*, que permite realizar convoluções de forma eficiente, proporcionando ao utilizador a possibilidade de auralizar uma sala em tempo real.

O trabalho aqui apresentado, e, em particular, o estudo de caso explorado nos capítulos 4 e 5, mostram que a auralização em tempo real com ouvinte móvel (*walkthrough auralisation*) baseada em modelos físicos é exequível desde que sejam pré-calculadas as respostas impulsivas relativas ao trajeto do ouvinte que se pretende simular. A interpolação espacial de respostas impulsivas é muito útil na medida em que pode diminuir o número de respostas impulsivas que é necessário obter por modelação, aliviando assim significativamente o peso computacional do processo. Porém, a interpolação linear de respostas impulsivas revelou-se insuficiente uma vez que não reflete a desejável interpolação de atrasos temporais. Foi por isso implementada interpolação em fase e módulo nos moldes sugeridos em [7]. Constatou-se uma dificuldade prática na utilização de BRIR interpoladas desta forma, traduzida no aparecimento de um eco perfeitamente audível. Não tendo sido encontrada na literatura nenhuma menção a este artefacto, foram desenvolvidos métodos empíricos de o resolver, com resultados satisfatórios quer do ponto de vista perceptual (o artefacto é suprimido) quer do ponto de vista do peso computacional (impacto praticamente nulo). Estes métodos contemplam o balanceamento da energia do sinal sintetizado relativamente aos sinais interpolados. São disponibilizados, nos anexos correspondentes deste documento, exemplos demonstrativos dos resultados obtidos.

### 6.2 Trabalho Futuro

#### 6.2.1 Perspectiva Geral

A experiência de estímulos sensoriais de forma artificial é o principal objeto de estudo da *Realidade Virtual* - *RV*. Sofisticados simuladores de voo, jogos de computador e cinema são

exemplos de aplicação prática onde a *RV* desempenha um papel preponderante. Os estímulos sonoros proporcionam, ao ser humano, percepção espacial sobre o meio que o rodeia. Neste sentido, a auralização é um fragmento da *RV* de extrema importância. O trabalho apresentado pretende constituir um contributo nessa área.

Nesta secção são apresentadas algumas propostas de trabalho futuro no sentido de melhorar as ferramentas desenvolvidas, quer do ponto de vista de modelação acústica que do ponto de vista da auralização em tempo real.

## 6.2.2 Modelação Acústica

- **Validação** - Com vista à validação do *DWsim* relativamente ao seu rigor relacionado com a absorção, reflexão e difração deverá ser realizada uma medição *in situ* de uma RIR mais os parâmetros RT, C80, D50 ou apenas pressão sonora (campo direto e reverberante) comparando os valores obtidos com os da norma.
- **Condições-Fronteira** - O *DWsim* considera um coeficiente de reflexão fixo para cada material. Como já referido anteriormente, o coeficiente de reflexão dos materiais depende da frequência de amostragem da onda incidente, bem como do ângulo de incidência. A aplicação de filtros FIR nos nós-fronteira pode contribuir para tornar o modelo mais realista. Mais informação sobre esta matéria pode ser encontrada em [8].
- **Topologias de Modelação** - Foi implementada a topologia retilínea tridimensional no simulador *DWsim*. Este poderia suportar outras topologias (por exemplo octaédrica) tornando-se assim mais versátil.
- **Paralelização** - É possível dividir a malha de nós que forma o modelo de guias-de-onda digitais e processar cada uma das partes em diferentes máquinas. A paralelização do processo de modelação permite diminuir o tempo de computação. Esta funcionalidade pode ser implementada no *DWsim*. A paralelização é um tema abordado em [9].
- **Integração com Plataformas de Desenho 3D** - A modelação acústica de salas constitui uma ferramenta muito importante para perceber quais as características que favorecem o campo sonoro de uma determinada sala que se pretende construir. Obter a matriz de pontos de um espaço físico desenhado numa plataforma do tipo *Google SketchUp* possibilitaria modelar acusticamente esse mesmo espaço virtual de uma forma relativamente simples recorrendo ao *DWsim*. O trabalho desenvolvido em [5] pode ser visto como um ponto de partida para atingir este objetivo.
- **Base de Dados de Materiais** - Pode ser criada uma base de dados que contenha os coeficientes de reflexão para diversos materiais. Assim, a mesma base de dados poderia ser usada por todos os utilizadores do *DWsim*.

- **Interface Gráfico** - Pode ser desenvolvido um ambiente gráfico que torne a utilização do *DWsim* mais fácil e agradável. O utilizador poderia visualizar o espaço físico e colocar a(s) fonte(s) e o(s) ouvinte(s) nas posições pretendidas.

### 6.2.3 Auralização em Tempo Real

- **Interpolação em Fase e Módulo** - Podem existir alternativas de interpolação que, mantendo os benefícios da interpolação linear de fase, garantam à partida a inexistência de eco (bem como de outros artefactos indesejáveis). Nesse sentido, parece útil investigar a possibilidade de obter analiticamente a expressão no domínio do tempo (iDFT) da resposta interpolada em fase e módulo.
- **Head-Tracking** - A utilização de sistemas de *tracking* confere maior liberdade de movimentos ao utilizador. Em vez de utilizar o teclado do computador para mover o ouvinte, poderia movimentar-se livremente num espaço em que a sua posição relativa determina a posição do ouvinte dentro da sala.
- **Visualização** - Além da auralização, o utilizador poderia usar simultaneamente dispositivos de visualização (*Head-Mounted Display* - *HMD*).
- **Interface Gráfico** - O desenvolvimento de um ambiente de interação gráfica em que o utilizador pode escolher facilmente a sala que pretende auralizar, e até definir um trajeto dentro da mesma, é uma mais-valia para a aplicação de auralização. A biblioteca de software desenvolvida - *OLAlib* - permite uma fácil integração com um sistema deste tipo.

## Apêndice A

# Caraterísticas Acústicas de Materiais

### A.1 Impedância Acústica Específica

Material	$v_l$	$\rho$	$Z$
Alumina	10520	3.86	40.6
Aluminium rolled	6420	2.70	17.33
Beryllium	12890	1.87	24.1
Bismuth	2200	9.8	21.5
Brick	4300	1.7	7.4
Cadmium	2800	8.6	24.0
Concrete	3100	2.6	8.0
Copper rolled	5010	8.93	44.6
Fused silica	5700	2.20	12.55
Germanium	5410	5.47	29.6
Glass pyrex	5640	2.24	13.1
Glass quartz	5500	2.2	12.1
Glass silica	5900	2.2	13.0
Glucose	3200	1.56	5.0
Gold	3240	19.7	63.8
Induim	2560	7.3	18.7
Iron	5900	7.69	46.4
Iron cast	4600	7.22	33.2
Lead	2200	11.2	24.6
Lithium	7080	4.7	33.0
Magnesium	5800	1.73	10.0
Marble	3800	2.8	10.5
Molybdenum	6300	10.0	63.1
Nickel	5600	8.84	49.5
Paraffin	1940	0.91	1.76
Polyester casting resin	2290	1.07	2.86
Porcelain	5900	2.3	13.5

Material	$v_l$	$\rho$	$Z$
Silicon carbide	13006	13.8	91.8
Silicon nitride	11000	3.27	36.0
Silver	3600	10.6	38.0
Steel mild	5900	7.8	46.0
Steel stainless	5790	7.89	45.7
Stycast	2220	1.19	2.64
Tantalum	4100	16.6	54.8
Teflon	1390	2.14	2.97
Tin	3300	7.3	24.2
Titanium	6100	4.48	27.3
Tracon	2970	1.62	4.82
Tungsten	5200	19.4	101.0
Uranium	3400	18.5	63.0
Vanadium	6000	6.03	36.2

*Tabela A.1: Impedância acústica específica de alguns materiais [10]*

---

$v_l = \text{longitudinal sound velocity [m/s]}$

$\rho = \text{density [g/cm}^3\text{]}$

$Z = \text{acoustic impedance [Rayl} \cdot 10^6\text{]}$

$$1\text{Rayl} = \frac{Kg}{s \cdot m^2}$$

## A.2 Coeficiente de Reflexão

Floor Materials	125 Hz	250 Hz	500 Hz	1 KHz	2 KHz	4 KHz
Carpet	0.99	0.98	0.94	0.8	0.75	0.55
Concrete (unpainted, rough finish)	0.99	0.98	0.96	0.94	0.92	0.90
Concrete (sealed or painted)	0.99	0.99	0.98	0.98	0.98	0.98
Marble or glazed tile	0.99	0.99	0.99	0.99	0.98	0.98
Vinyl tile or linoleum on concrete	0.98	0.97	0.97	0.97	0.97	0.98
Wood parquet on concrete	0.96	0.96	0.93	0.94	0.94	0.93
Wood flooring on joists	0.85	0.89	0.9	0.93	0.94	0.93
Seating Materials	125 Hz	250 Hz	500 Hz	1 KHz	2 KHz	4 KHz
Benches (wooden, empty)	0.9	0.91	0.92	0.92	0.92	0.92
Benches (wooden, 2/3 occupied)	0.63	0.6	0.53	0.47	0.44	0.44
Benches (wooden, fully occupied)	0.5	0.44	0.34	0.24	0.2	0.24
Benches (cushioned seats and backs, empty)	0.68	0.6	0.58	0.56	0.457	0.52
Benches (cushioned seats and backs, fully occupied)	0.56	0.44	0.34	0.28	0.28	0.33
Benches (cushioned seats and backs, 2/3 occupied))	0.5	0.36	0.24	0.14	0.14	0.24
Theater seats (wood, empty)	0.97	0.96	0.95	0.93	0.92	0.92
Theater seats (wood, 2/3 occupied)	0.66	0.79	0.72	0.43	0.44	0.44
Theater seats (wood, fully occupied)	0.5	0.7	0.6	0.24	0.20	0.24
Seats (fabric-upholsterd, empty)	0.51	0.44	0.2	0.12	0.18	0.3
Seats (fabric-upholsterd, fully occupied)	0.4	0.26	0.12	0.04	0.07	0.15



Reflective wall materials	125 Hz	250 Hz	500 Hz	1 KHz	2 KHz	4 KHz
Brick (natural)	0.97	0.97	0.97	0.96	0.95	0.93
Brick (painted)	0.99	0.99	0.98	0.98	0.98	0.97
Concrete block (coarse)	0.64	0.56	0.69	0.71	0.61	0.75
Concrete block (painted)	0.90	0.95	0.94	0.93	0.91	0.92
Concrete (poured, rough finish, unpainted)	0.99	0.98	0.96	0.94	0.92	0.90
Doors (solid wood panels)	0.90	0.93	0.95	0.96	0.96	0.96
Glass (1/4"plate, large pane)	0.82	0.94	0.96	0.97	0.98	0.98
Glass (small pane)	0.96	0.96	0.97	0.97	0.98	0.98
Plasterboard (12mm (1/2") paneling on studs)	0.71	0.90	0.94	0.95	0.96	0.96
Plaster (gypsum or lime, on masonry)	0.99	0.98	0.98	0.97	0.96	0.95
Plaster (gypsum or lime, on wood lath)	0.86	0.90	0.94	0.95	0.96	0.96
Ceiling Material	125 Hz	250 Hz	500 Hz	1 KHz	2 KHz	4 KHz
Plasterboard (12mm(1/2") in suspended ceiling grid)	0.85	0.89	0.96	0.96	0.93	0.92
Underlay in perforated metal panels (25mm(1") batts)	0.49	0.22	0.43	0.23	0.10	0.21
Metal deck (perforated channels,25mm(1") batts)	0.81	0.31	0.01	0.12	0.48	0.73
Metal deck (perforated channels, 75mm(3") batts)	0.27	0.01	0.01	0.11	0.48	0.69
Plaster (gypsum or lime, on masonry)	0.99	0.98	0.98	0.97	0.96	0.95
Sprayed cellulose fiber (16mm(5/8") on solid backing)	0.95	0.84	0.56	0.21	0.10	0.09
Sprayed cellulose fiber (25mm(1") on solid backing)	0.92	0.71	0.25	0.02	0.07	0.24
Sprayed cellulose fiber (32mm(1-1/4") on solid backing)	0.90	0.70	0.27	0.08	0.02	0.02
Sprayed cellulose fiber (75mm(3") on solid backing)	0.30	0.05	0	0.15	0.15	0.10
Wood tongue-and-groove roof decking	0.76	0.81	0.86	0.92	0.87	0.90
Miscellaneous surface material	125 Hz	250 Hz	500 Hz	1 KHz	2 KHz	4 KHz
People-adults (per 1/10 person)	0.75	0.65	0.58	0.54	0.50	0.50
People-high school students (per 1/10 person)	0.78	0.70	0.62	0.58	0.55	0.55
People-elementary students (per 1/10 person)	0.82	0.77	0.72	0.68	0.65	0.65
Ventilating grilles	0.70	0.60	0.50	0.50	0.50	0.60
Water or ice surface	0.99	0.99	0.98	0.98	0.98	0.97

Tabela A.2: Coeficientes de reflexão de alguns materiais comuns [11]

## Apêndice B

# Código-Fonte do Software de Modelação *DWsim*

Este anexo remete para a pasta 'AnexoB' contida no *CD* que se encontra na contra-capa deste documento. Contém os ficheiros relativos ao código fonte (*C++*) do *DWsim* apresentado no capítulo 3.

## Apêndice C

# Código-Fonte da Biblioteca *OLAlib*

Este anexo remete para a pasta 'AnexoC' contida no *CD* que se encontra na contra-capa deste documento. Contém os ficheiros relativos ao código fonte (*Matlab*) da biblioteca *OLAlib* apresentada no capítulo 4.

## Apêndice D

# Auralização num Trajeto Virtual

Este anexo remete para a pasta 'AnexoD' contida no *CD* que se encontra na contra-capa deste documento. Contém ficheiros áudio relativos à auralização da sala utilizada no estudo de caso apresentado no capítulo 4. O ouvinte desloca-se desde a posição inicial (no corredor) até à posição final a uma velocidade de  $30\text{cm/s}$ . É utilizada a resposta impulsional bi-aural (BRIR) mais próxima para cada posição do ouvinte.

Foram utilizados diferentes sinais de entrada no processo de auralização. A tabela seguinte apresenta o nome dos ficheiros em causa e o respetivo sinal de entrada.

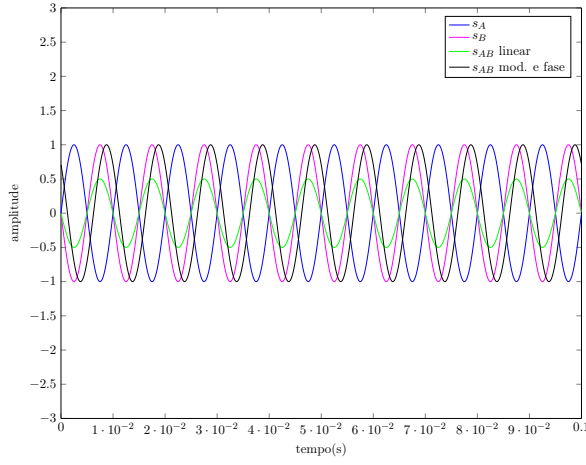
Ficheiro de Auralização	Sinal de Entrada
path_speech.wav	discurso
path_sin440.wav	sinusóide a 440 Hz
path_sin1k.wav	sinusóide a 1KHz
path_sin5k.wav	sinusóide a 5KHz

## Apêndice E

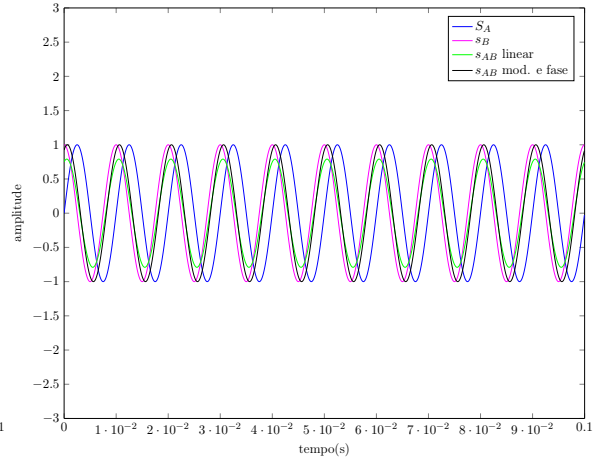
# Interpolação de Sinusóides

Neste anexo são apresentados alguns resultados da interpolação entre duas sinusóides. Os factores de ponderação da interpolação,  $\alpha$  e  $\beta$ , são indicados em cada figura bem como o desfasamento,  $\phi$  (em radianos), entre as sinusóides. As sinusóides intervenientes na operação de interpolação são representadas por  $s_A$  e  $s_B$ . A sinusóide resultante é representada por  $s_{AB}$ . Em cada gráfico observa-se o resultado da interpolação linear e o resultado da interpolação em módulo e fase.

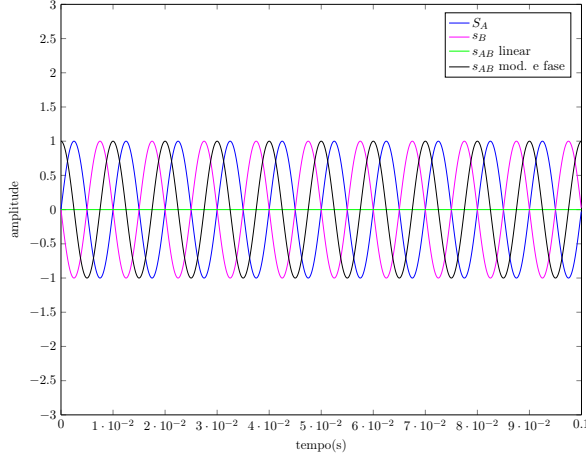
Relembre-se que  $\alpha + \beta = 1$ .



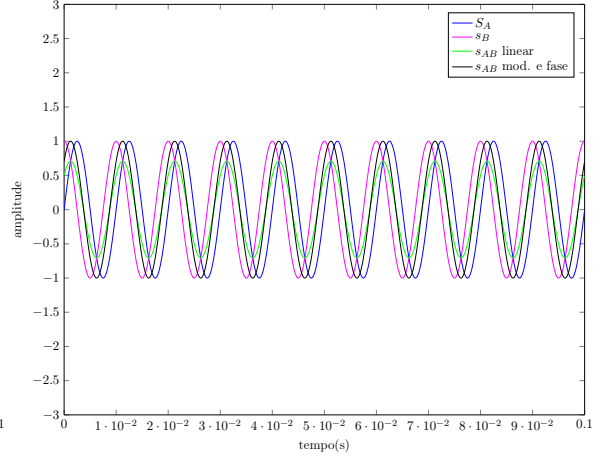
(a)  $\alpha = 0.25, \beta = 0.75, \phi = \pi$



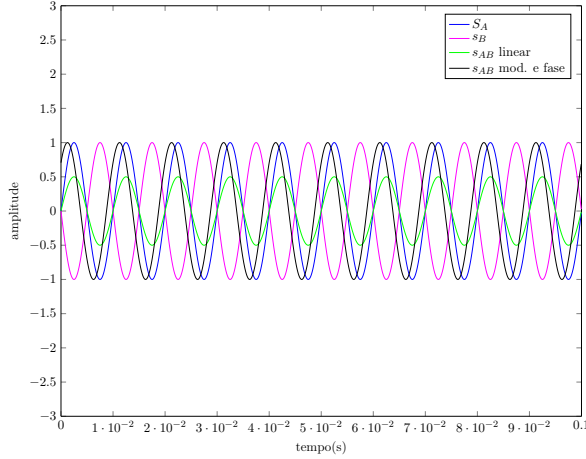
(b)  $\alpha = 0.25, \beta = 0.75, \phi = \frac{\pi}{2}$



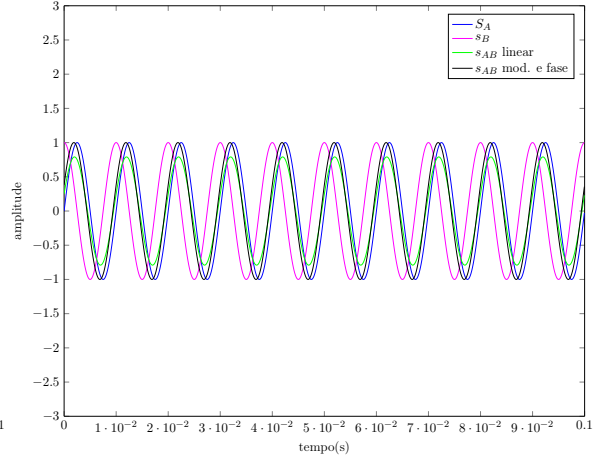
(c)  $\alpha = 0.50, \beta = 0.50, \phi = \pi$



(d)  $\alpha = 0.50, \beta = 0.50, \phi = \frac{\pi}{2}$

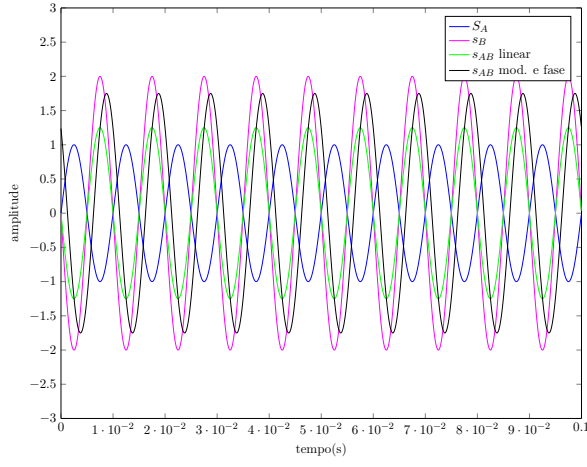


(e)  $\alpha = 0.75, \beta = 0.25, \phi = \pi$

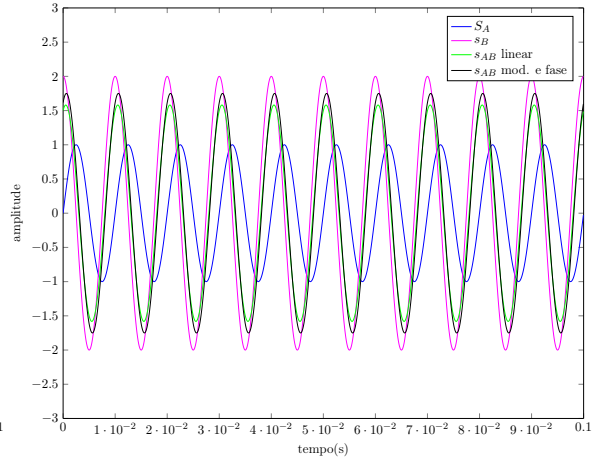


(f)  $\alpha = 0.75, \beta = 0.25, \phi = \frac{\pi}{2}$

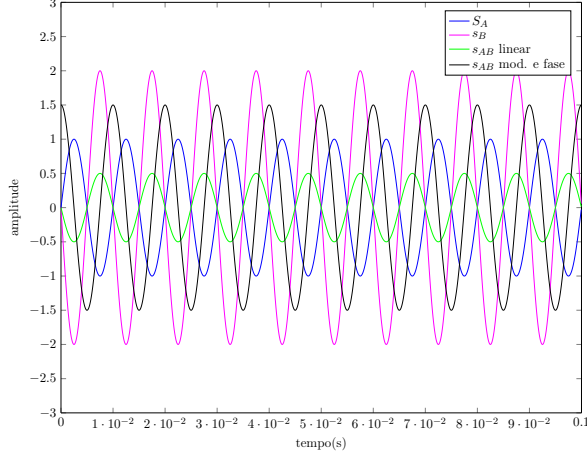
Figura E.1: Resultados da interpolação entre sinusóides de igual amplitude



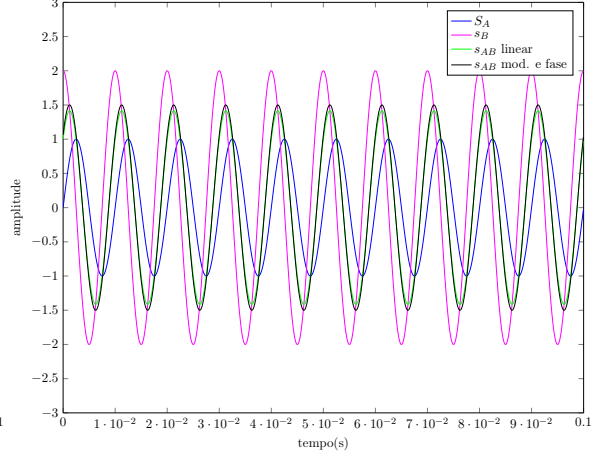
(a)  $\alpha = 0.25, \beta = 0.75, \phi = \pi$



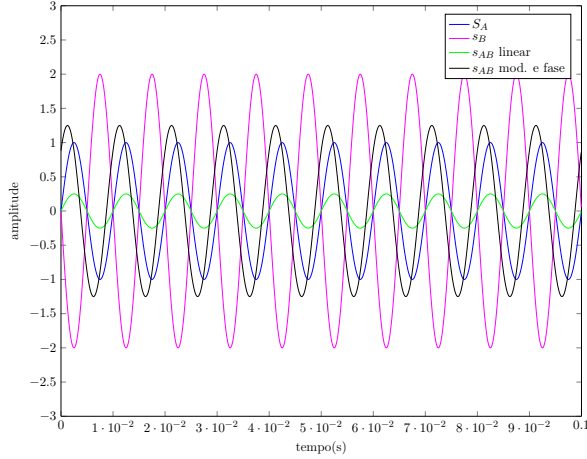
(b)  $\alpha = 0.25, \beta = 0.75, \phi = \frac{\pi}{2}$



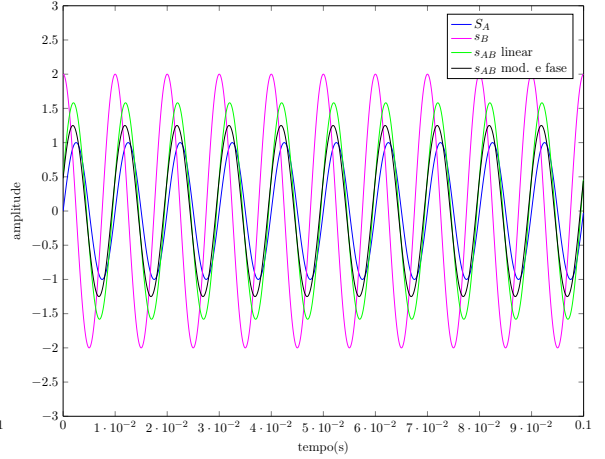
(c)  $\alpha = 0.50, \beta = 0.50, \phi = \pi$



(d)  $\alpha = 0.50, \beta = 0.50, \phi = \frac{\pi}{2}$

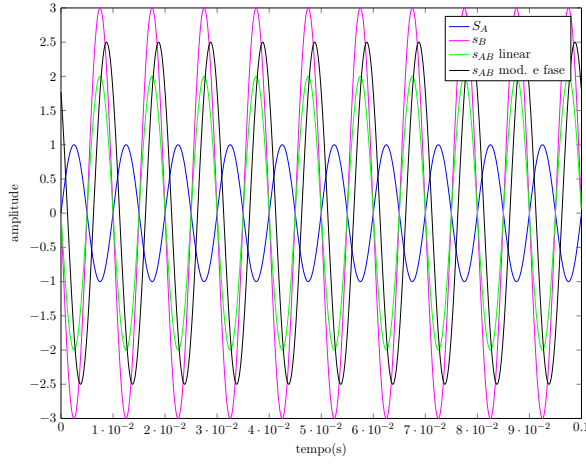


(e)  $\alpha = 0.75, \beta = 0.25, \phi = \pi$

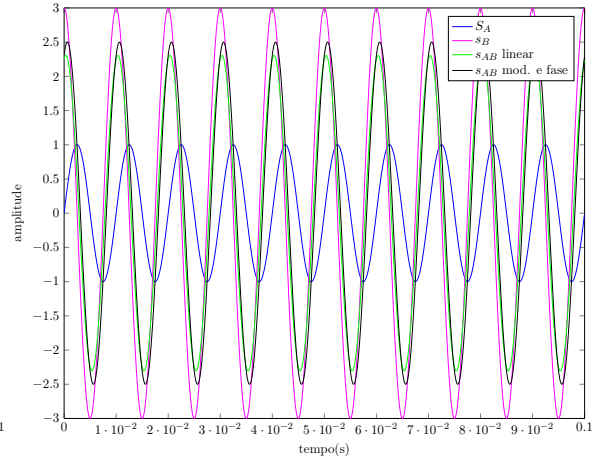


(f)  $\alpha = 0.75, \beta = 0.25, \phi = \frac{\pi}{2}$

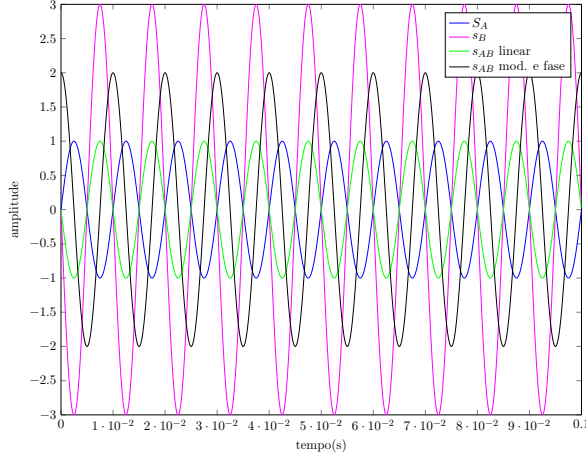
Figura E.2: Resultados da interpolação entre sinusóides em que uma tem o dobro da amplitude



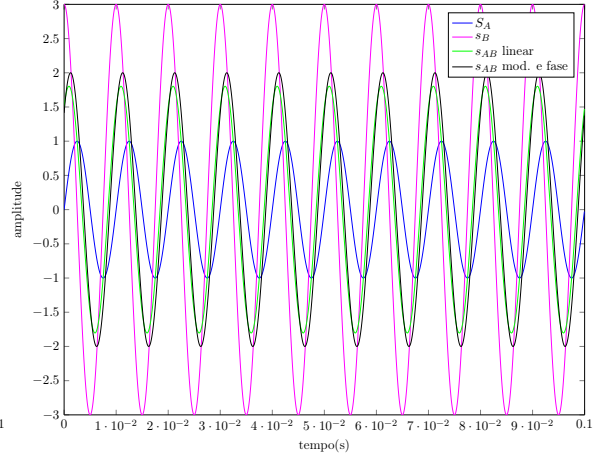
(a)  $\alpha = 0.25, \beta = 0.75, \phi = \pi$



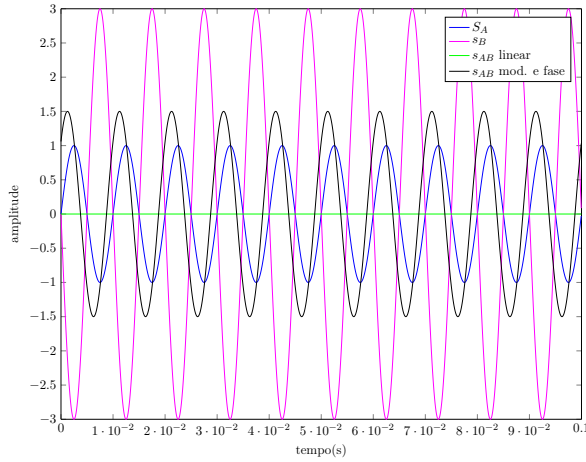
(b)  $\alpha = 0.25, \beta = 0.75, \phi = \frac{\pi}{2}$



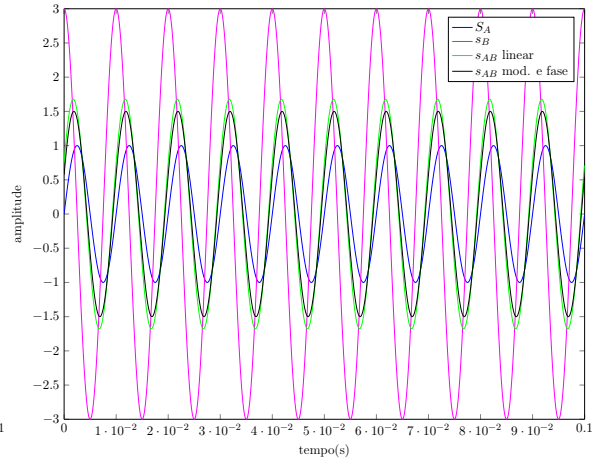
(c)  $\alpha = 0.50, \beta = 0.50, \phi = \pi$



(d)  $\alpha = 0.50, \beta = 0.50, \phi = \frac{\pi}{2}$



(e)  $\alpha = 0.75, \beta = 0.25, \phi = \pi$



(f)  $\alpha = 0.75, \beta = 0.25, \phi = \frac{\pi}{2}$

Figura E.3: Resultados da interpolação entre sinusóides em que uma tem o triplo da amplitude



## Apêndice F

# Interpolação em Módulo e Fase de uma RIR

Este anexo remete para a pasta 'AnexoF' contida no *CD* que se encontra na contra-capas deste documento. Contém o ficheiro áudio resultante da interpolação em fase e módulo entre duas respostas impulsivas. Cada resposta impulsional tem duração de um segundo a uma frequência de amostragem de  $44100\text{ Hz}$  - 44100 amostras.

Ficheiro de Áudio	Descrição
anech_speech.wav	áudio anecóico
interpol_md_speech_eco1.wav	resposta interpolada: eco 1s
interpol_md_speech_eco2.wav	resposta interpolada: eco 2s
interpol_md_speech_cut.wav	resposta interpolada: estendida ( <i>zero-padding</i> ) e cortada

## Apêndice G

# Resultados dos Testes de Interpolação

Este anexo remete para a pasta 'AnexoG' contida no *CD* que se encontra na contra-capa deste documento. Contém os ficheiros áudio correspondentes aos testes realizados segundo o esquema apresentado no capítulo 5. São apresentados resultados relativos à interpolação linear e testes relativos à interpolação em módulo e fase. As seguintes tabelas contêm o nome dos ficheiros áudio e a respectiva descrição.

Ficheiro de Áudio	Descrição
ha.wav	auralização na posição $h_A$
hb.wav	auralização na posição $h_B$
hc.wav	auralização na posição $h_C$
hd.wav	auralização na posição $h_C$
he.wav	auralização na posição $h_C$

*Tabela G.1: Auralização das respostas impulsiais simuladas na posições assinaladas no esquema de testes apresentado no capítulo 5*

Ficheiro de Áudio	Descrição
ha_he_lin_speech.wav	interpolação entre $h_A$ e $h_E$ (comparar com $h_B$ )
hb_he_lin_speech.wav	interpolação entre $h_B$ e $h_E$ (comparar com $h_C$ )
hc_he_lin_speech.wav	interpolação entre $h_C$ e $h_E$ (comparar com $h_D$ )

*Tabela G.2: Resultados da interpolação linear entre as respostas impulsiais indicadas. O sinal áudio de entrada é um sinal de voz (discurso)*

Ficheiro de Áudio	Descrição
ha_he_lin_sin440.wav	interpolação entre $h_A$ e $h_E$ (comparar com $h_B$ )
hb_he_lin_sin440.wav	interpolação entre $h_B$ e $h_E$ (comparar com $h_C$ )
hc_he_lin_sin440.wav	interpolação entre $h_C$ e $h_E$ (comparar com $h_D$ )

*Tabela G.3: Resultados da interpolação linear entre as respostas impulsiais indicadas. O sinal áudio de entrada é uma sinusóide a 440Hz*

Ficheiro de Áudio	Descrição
ha_he_md_speech.wav	interpolação entre $h_A$ e $h_E$ (comparar com $h_B$ )
hb_he_md_speech.wav	interpolação entre $h_B$ e $h_E$ (comparar com $h_C$ )
hc_he_md_speech.wav	interpolação entre $h_C$ e $h_E$ (comparar com $h_D$ )

*Tabela G.4: Resultados da interpolação em módulo e fase entre as respostas impulsiais indicadas. O sinal áudio de entrada é um sinal de voz (discurso)*

Ficheiro de Áudio	Descrição
ha_he_md_sin440.wav	interpolação entre $h_A$ e $h_E$ (comparar com $h_B$ )
hb_he_md_sin440.wav	interpolação entre $h_B$ e $h_E$ (comparar com $h_C$ )
hc_he_md_sin440.wav	interpolação entre $h_C$ e $h_E$ (comparar com $h_D$ )

*Tabela G.5: Resultados da interpolação em módulo e fase entre as respostas impulsiais indicadas. O sinal áudio de entrada é uma sinusóide a 440Hz*



# Bibliografia

- [1] Damian Murphy, Antti Kelloniemi, Jack Mullen, and Simon Shelley. Acoustic Modeling Using the Digital Waveguide Mesh. *IEEE, Signal Processing Magazine*, pg. 55-66, Vol 24-2, 2007.
- [2] L Savioja, T Lokki, and V Valimäki. The interpolated 3-D digital waveguide mesh method for room acoustic simulation and auralization. *Joint Baltic-Nordic Acoustical Meeting*, 2002.
- [3] A Southern, D Murphy, T Lokki, and L Savioja. The Perceptual Effects of Dispersion Error on Room Acoustic Model Auralization. *Proc. Forum Acusticum*, 2011.
- [4] Tom Barker, Guilherme Campos, Paulo Dias, José Vieira, Catarina Mendonça, and Jorge A. Santos. Real-time auralisation system for virtual microphone positioning. *15th Int. Conference on Digital Audio Effects (DAFx-12)*, 2012.
- [5] Jorge Miguel de Almeida Pereira. Aquisição e tratamento de dados 3D para modelação acústica de salas. MSc, UA. 2013.
- [6] Gavin Kearney, Claire Masterson, Stephen Adams, and Frank Boland. Dynamic Time Warping for Acoustic Response Interpolation : Possibilities and Limitations. 2009.
- [7] Sanjeev Mehrotra, Wei-ge Chen, and Zhengyou Zhang. Interpolation of combined head and room impulse response for audio spatialization. *2011 IEEE 13th International Workshop on Multimedia Signal Processing*, 2011.
- [8] Antti Kelloniemi, Lauri Savioja, Vesa Välimäki, and Senior Member. Spatial Filter-Based Absorbing Boundary for the 2-D Digital Waveguide Mesh. 2005.
- [9] António Guilherme Rocha Campos. *Three-Dimensional Digital Waveguide Mesh Modeling for Room Acoustics Simulation*. PhD thesis, University of York, 2003.
- [10] Acoustic Impedance of Some Solids. [http://traktoria.org/files/sonar/passive\\_materials/acoustic\\_impedance\\_of\\_some\\_solids.htm](http://traktoria.org/files/sonar/passive_materials/acoustic_impedance_of_some_solids.htm). Accessed: 2015-12-4.
- [11] Absorption Coefficients of Common Building Materials and Finishes. <http://www.soundproofyourhome.com/absorption-coefficient-chart>. Accessed: 2015-12-4.
- Alex Southern, Jeremy Wells, and Damian Murphy. Rendering Walk-Through Auralisations Using Wave-Based Acoustical Models. *17th European Signal Processing Conference*, 2009.

- A Kelloniemi. Improved adjustable boundary condition for the 3-D digital waveguide mesh. *Applications of Signal Processing to Audio and Acoustics, 2005. IEEE Workshop on*, 2005.
- D T Murphy and J Mullen. Digital waveguide mesh modelling of room acoustics: improved anechoic boundaries. *Proc. DAFX-02*, 2002.
- A. Oliveira and G. Campos. Real-time dynamic image-source implementation for auralisation. *Proceedings of the 16th International Conference on Digital Audio Effects. Maynooth, 2013. p. 368-372. Dafx*, 2013.
- Rémi Mignot, Laurent Daudet, and François Ollivier. Room Reverberation Reconstruction : Interpolation of the Early Part Using Compressed Sensing. 2013.
- Antti Kelloniemi. *Room Acoustics Modeling With the Digital Waveguide Mesh* . 2006.
- Alex Southern and Damian Murphy. First Order Source Directivity Excitation for Wave-Based Room Acoustics Simulation. *EAA Symposium on Auralization*, 2009.
- Damian T Murphy, Chris J C Newton, and David M Howard. Digital Waveguide Mesh Modelling Of Room Acoustics: Surround-Sound, Boundaries and Plugin Implentation. *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-01)*, 2001.
- Yiu Wai Lam. Issues for computer modelling of room acoustics in non-concert hall settings. 2005.